Copy for the Elected Office (EO/US) PCT/ES99/00017

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and Administrative Instructions, Section 422)

To:

UNGRÍA LÓPEZ, Javier
Avda. Ramón y Cajal, 78
E-28043 Madrid
ESPAGNE

Date of mailing (day/month/year)
02 December 1999 (02.12.99)

Applicant's or agent's file reference
199.171/MAD

IMPORTANT NOTIFICATION

International application No.
PCT/ES99/00017

International filing date (day/month/year)
23 January 1999 (23.01.99)

1. The following indications appeared on record concerning:

☐ the applicant ☐ the inventor ☒ the agent ☐ the common representative

Name and Address

OJEDA GARCIA, Pedro
Consejo Superior de Investigaciones Científicas
Calle Serrano, 113
E-28006 Madrid
Spain

State of Nationality:
State of Residence:
Telephone No.: 91 585 52 76
Facsimile No.: 91 585 52 87
Teleprinter No.:

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☒ the person ☐ the name ☐ the address ☐ the nationality ☐ the residence

Name and Address

UNGRÍA LÓPEZ, Javier
Avda. Ramón y Cajal, 78
E-28043 Madrid
Spain

State of Nationality:
State of Residence:
Telephone No.: 34 91 413 60 62
Facsimile No.: 34 91 413 64 17
Teleprinter No.:

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

☒ the receiving Office
☐ the International Searching Authority
☒ the International Preliminary Examining Authority
☐ the designated Offices concerned
☒ the elected Offices concerned
☒ other: OJEDA GARCIA, Pedro

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer
F. Gateau

Telephone No.: (41-22) 338.83.38

Form PCT/IB/306 (March 1994)

002985563

PATENT COOPERATION TREAT.

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Assistant Commissioner for Patents
United States Patent and Trademark Office
Box PCT
Washington, D.C.20231
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 13 October 1999 (13.10.99)	
International application No. PCT/ES99/00017	**Applicant's or agent's file reference**
International filing date (day/month/year) 23 January 1999 (23.01.99)	**Priority date** (day/month/year) 23 January 1998 (23.01.98)
Applicant PRIETO-DAPENA, Maria Pilar et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

12 August 1999 (12.08.99)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland**	Authorized officer Philippe Bécamel
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

2892892

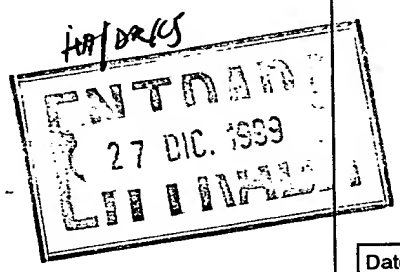
Form PCT/IB/331 (July 1992)

PATENT COOPERATION TREATY

From the:
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

UNGRIA LOPEZ, Javier
Avda. Ramon y Cajal, 78
28043 Madrid
ESPAGNE



PCT

WRITTEN OPINION

(PCT Rule 66)

Date of mailing *(day/month/year)*	20. 12. 99
Applicant's or agent's file reference: 199.171/MAD.	**REPLY DUE** **within 3 month(s)** from the above date of mailing

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/ES99/00017	23/01/1999	23/01/1998

International Patent Classification (IPC) or both national classification and IPC

C12N15/82

Applicant

CONSEJO SUPERIOR DE INVESTIGACIONES... et al.

1. This written opinion is the **first** drawn up by this International Preliminary Examining Authority.

2. This opinion contains indications relating to the following items:

I	☒	Basis of the opinion
II	☐	Priority
III	☐	Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
IV	☐	Lack of unity of invention
V	☒	Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
VI	☐	Certain document cited
VII	☐	Certain defects in the international application
VIII	☒	Certain observations on the international application

3. The applicant is hereby **invited to reply** to this opinion.

When? See the time limit indicated above. The applicant may, before the expiration of that time limit, request this Authority to grant an extension, see Rule 66.2(d).

How? By submitting a written reply, accompanied, where appropriate, by amendments, according to Rule 66.3. For the form and the language of the amendments, see Rules 66.8 and 66.9.

Also: For an additional opportunity to submit amendments, see Rule 66.4.
For the examiner's obligation to consider amendments and/or arguments, see Rule 66.4 bis.
For an informal communication with the examiner, see Rule 66.6.

If no reply is filed, the international preliminary examination report will be established on the basis of this opinion.

4. The final date by which the international preliminary examination report must be established according to Rule 69.2 is: 23/05/2000.

Name and mailing address of the international preliminary examining authority:
European Patent Office
D-80298 Munich
Tel. +49 89 2399 - 0 Tx: 523656 epmu d
Fax: +49 89 2399 - 4465

Authorized officer / Examiner
Burkhardt, P

Formalities officer (incl. extension of time limits)
Vullo, C
Telephone No. +49 89 2399 8061

I. Basis of the opinion

1. This opinion has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this opinion as "originally filed".*):

Description, pages:

1-23 as originally filed

Claims, No.:

1,2,5-25 as received on 24/08/1999 with letter of 19/08/1999

Drawings, sheets:

1/5-5/5 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. This opinion has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

see separate sheet

4. Additional observations, if necessary:

see separate sheet

V. Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims	
Inventive step (IS)	Claims	1, 2, 5 - 25 (NO)
Industrial applicability (IA)	Claims	

2. Citations and explanations

s e s parat sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

see separate sheet

Re Item I
Basis of the opinion

The amended claims 1, 2 and 5 - 25 filed with the letter of 19.08.1999 are formally acceptable under Articles 19(2) and 34(2)(b) PCT.
Amended claims 3 and 4 are not acceptable under Articles 19(2) and 34(2)(b) PCT. They read on to any nucleotide sequence comprising small fragments or short nucleotide sequences of SEQ ID NO:1 and therefore the amendments go beyond the disclosure in the international application as filed (Articles 19(2) and 34(2)(b) PCT).

This written opinion is therefore based on amended claims 1, 2 and 5 - 25 (Rule 70.2(c) PCT).

This written opinion is also based on the Sequence Listing (pages 1-2) as filed with the letter of 19.08.1999.

Re Item V
Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

The following document (D) is referred to in this communication; the numbering will be adhered to in the rest of the procedure:

D1 Almoguera and Jordano, 1992. Plant Mol. Biol. 19:781-792.

1. Article 33(2) PCT (Novelty)

1.1 For the interpretation of the present set of claims please see section VIII.

1.2 Present claim 1 is directed to the **genomic** sequence of the sunflower *Ha ds10 G1* gene (SEQ ID NO:1). Prior art document D1 discloses the **cDNA** sequence of the sunflower *Ha ds10 G1* gene. Thus, D1 does not anticipate the subject-matter of present claim 1. The same holds true for dependent claim 2 and

for present claims 5 - 17 addressing homologous sequences, expression cassettes, vectors and host cells containing said sequence.

1.3 The use of *Ha ds10 G1* sequences for seed- or seedling-specific expression of chimeric genes in transgenic plants, as laid out in present claims 18 - 20 as well as the resulting plants and the use of these plants (claims 21 - 24) has not yet been disclosed in the prior art presently available to the IPEA.

1.4 For the assessment of novelty of the present "product by process" claim 25 no unified criteria exist in the PCT. The EPO, for example, does not recognize novelty merely by the fact that the product is produced by means of a new process. Novelty can only be established where use of the method necessarily means that the product has a particular characteristic and that a person skilled in the art following the teaching of the application would inevitably acquire a product which has different characteristics to the product disclosed in the prior art. This does not seem to be the case for present claim 25.

2. Article 33(3) PCT (Inventive step)

2.1 The closest prior art to the subject-matter of present claim 1 appears to be D1. It discloses the cDNA sequence of the sunflower *Ha ds10 G1* gene (page 785, Figure 1) and furthermore states that isolation and characterization of the corresponding genomic sequences will allow further studies on the regulation of the gene (D1, page 790, last paragraph). Claim 1 differs from that in the presentation of the genomic sequence of the sunflower *Ha ds10 G1* gene.

2.2 In the light of the prior art and having regard to the present description and claims, the technical problem may thus be the provision of the genomic *Ha ds10 G1* sequence.
It is common general knowledge in modern biotechnology, and therefore within the scope of a man skilled in the art, to isolate the genomic sequence of a gene for which the cDNA sequence is known. Therefore, the subject-matter of present claim 1 is not based on an inventive concept. The same holds true for present claims 2 and 5 - 17.

2.3 A similar objection applies to present claim 18 directed to the use of the above nucleotide sequences for specific expression of chimeric genes in seeds, seed parts, seed extract, seed embryo, and seedling tissue. It has been known from the prior art (D1, page 787, left column, second paragraph) that the transcript of the *Ha ds10 G1* especially accumulates in embryos, dry seeds and seedlings. Therefore, it appears to be obvious to use the corresponding promoter sequences to drive the expression of chimeric genes in the respective tissues. Present claim 18 does not meet the requirements of Article 33(3) PCT. The same holds true for present claims 19 - 25.

Re Item VIII

Certain observations on the international application

1. The use of internal arbitrary designations for the nucleotide sequence in claim 1 is meaningless to the person skilled in the art and does not constitute a definition through technical features as required by Rule 6.3 (a) PCT. A nucleotide sequence should be clearly and unambiguously characterized, e.g. by reference to a SEQ ID NO. In order to assist the applicant and to allow a meaningful examination claim 1 is interpreted as being directed to SEQ ID NO:1.

2. Present claim 1 is directed to a product ("nucleotide sequence") and a process ("use ... in ... gene expression"). The IPEA considers such a combination of claim categories as unclear (Article 6 PCT). Claim 1 should therefore be amended. The examination has been limited to the product part of claim 1.

17

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
199.171/MAD.		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/ES99/00017	23/01/1999	23/01/1998
International Patent Classification (IPC) or national classification and IPC C12N15/82		
Applicant CONSEJO SUPERIOR DE INVESTIGACIONES... et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 6 sheets, including this cover sheet.

   ☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 3 sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
19/08/1999	04.05.00
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Burkhardt, P Telephone No. +49 89 2399 7456

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/ES99/00017

I. Basis of th report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-23	as originally filed			

Claims, No.:

1-24	as received on	02/03/2000	with letter of	28/02/2000

Claims, pages:

28-30	as received on	02/03/2000	with letter of	28/02/2000

Drawings, sheets:

1/5-5/5	as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

see separate sheet

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims 1-24
	No:	Claims
Inventive step (IS)	Yes:	Claims
	No:	Claims 1-24
Industrial applicability (IA)	Yes:	Claims 1-24
	No:	Claims

2. Citations and explanations

see separate sheet

Re Item I

Basis of the opinion

The amended claims filed with the letter of 28.02.2000 are formally acceptable under Article 34(2)(b) PCT.

This written opinion is also based on the Sequence Listing (pages 1-2) as filed with the letter of 19.08.1999.

Re Item V

Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

The following document (D) is referred to in this report:

D1 Almoguera and Jordano, 1992. Plant Mol. Biol. 19:781-792.

1.Article 33(2) PCT (Novelty)

1.1 Present claim 1 is directed to the **genomic** sequence of the sunflower *Ha ds10 G1* gene (SEQ ID NO:1). Prior art document D1 discloses the **cDNA** sequence of the sunflower *Ha ds10 G1* gene. Thus, D1 does not anticipate the subject-matter of present claim 1. The same holds true for dependent claim 2 - 16 addressing homologous sequences, expression cassettes, vectors and host cells containing said sequence.

1.2 The use of *Ha ds10 G1* sequences for seed- or seedling-specific expression of chimeric genes in transgenic plants, as laid out in present claims 17 - 19 as well as the resulting plants and the use of these plants (claims 20 - 23) has not yet been disclosed in the prior art presently available to the IPEA.

1.3 For the assessment of novelty of the present "product by process" claim 24 no unified criteria exist in the PCT. The EPO, for example, does not recognize novelty merely by the fact that the product is produced by means of a new

process. Novelty can only be established where use of the method necessarily means that the product has a particular characteristic and that a person skilled in the art following the teaching of the application would inevitably acquire a product which has different characteristics to the product disclosed in the prior art. This does not seem to be the case for present claim 25.

2. Article 33(3) PCT (Inventive step)

2.1 The closest prior art to the subject-matter of present claim 1 appears to be D1. It discloses the cDNA sequence of the sunflower *Ha ds10 G1* gene (page 785, Figure 1) and furthermore states that isolation and characterization of the corresponding genomic sequences will allow further studies on the regulation of the gene (D1, page 790, last paragraph). Claim 1 differs from that in the presentation of the genomic sequence of the sunflower *Ha ds10 G1* gene.

2.2 In the light of the prior art and having regard to the present description and claims, the technical problem may thus be the provision of the genomic *Ha ds10 G1* sequence.
It is common general knowledge in modern biotechnology, and therefore within the scope of a man skilled in the art, to isolate the genomic sequence of a gene for which the cDNA sequence is known. Therefore, the subject-matter of present claim 1 is not based on an inventive concept. The same holds true for present claims 2- 16.

2.3 The expression pattern of the *ds10* gene may be different from that of other *lea* and *lea-a* genes, although the *ds10* expression pattern itself was already known from the prior art (D1, page 787, left column, second paragraph).
Present claim 1, however, is directed to a product, i.e. the complete genomic *ds10* sequence. Its expression pattern is an inherent feature of the promoter sequence. With regard to the technical problem to be solved (see 2.2) this aspect therefore is neglectable.

2.4 A similar objection as in paragraph 2.2 applies to present claim 17 directed to the use of the above nucleotide sequences for specific expression of chimeric

genes in seeds, seed parts, seed extract, seed embryo, and seedling tissue. It has been known from the prior art (D1, page 787, left column, second paragraph) that the transcript of the *Ha ds10 G1* especially accumulates in embryos, dry seeds and seedlings. Therefore, it appears to be obvious to use the corresponding promoter sequences to drive the expression of chimeric genes in the respective tissues. Present claim 18 does not meet the requirements of Article 33(3) PCT. The same holds true for present claims 19 - 24.

CLAIMS

1. A nucleotide sequence constituted by the *Ha ds10 G1* gene, its promoter, Ha ds10 G1 5'- and 3' flanking sequences, wherein the nucleotide sequence is selected from the group consisting of identical nucleotide sequences identical to of SEQ ID NO:1, first homologous nucleotide sequences being homologous by at least 70% to SEQ ID NO:1, second homologous nucleotide sequences being homologous being at least 70% homologous to complementary sequences to SEQ ID NO:1, and fragments thereof.

2. A nucleotide sequence according to claim 1, wherein the first homologous sequence is homologous by at least 80% to SEQ ID NO:1.

3. A nucleotide sequence according to claim 1, wherein the first homologous sequence is homologous by less than 95% to SEQ ID NO:1.

4. A nucleotide sequence, wherein the second homologous sequence is homologous by at least 80% to SEQ ID NO:1.

5. A nucleotide sequence according to claim 1, wherein the second homologous sequence is homologous by less than 95% to SEQ ID NO:1.

6. A nucleotide sequence according to any of the claims 1 to 6, and further including a chimeric gene.

7. A nucleotide sequence according to claim 6, suitable for expression of a chimeric gene.

8. A nucleotide sequence according to claim 7, wherein the chimeric gene is specifical of seeds from early maturation stages.

9. A nucleotide sequence according to claim 8, constituted by constructions ds10F1, ds10F2, ds102Δ, ds10F3 and ds10EC1 or part thereof.

10. A nucleotide sequence according to claim 10, including *Ha ds10 G1* gene coding and 3'-flanking sequences.

11. A nucleotide sequence according to claim 10, including ds10F2 and ds10F2Δ in constructions.

12. A nucleotide sequence according to claim 8, including *Ha ds10 G1* gene coding and intron sequences.

13. A nucleotide sequence according to claim 12, contained in constructions ds10F3.

14. An expression cassette including a nucleotide sequence according to any of claims 1 to 13 and a chimeric gene.

15. A vector including an expression cassette according to claim 14.

16. Host cells including a nucleotide sequence according to any of claims 14 to 15.

17. Use of nucleotide sequences as defined in any of claims 1 to 15, in the specific expression of chimeric genes in seeds, seed parts, seed extract, seed embryos and seedling tissues.

18. Use of nucleotide sequences as defined in any of claims 9 to 11 for increasing the expression of chimeric genes specifically in transgenic plant seeds.

19. Use of nucleotide sequences as defined in any of claims 11 to 13 for increasing the expression of chimeric genes in seeds and/or reduce it in other tissues.

20. A transgenic plant transformed by a nucleotide sequence according to any of claims 1 to 15.

21. A transgenic plant according to claim 20, selected from sunflower, tobacco, soya, oilseed rape, "canola", maize, wheat, barley, rice, bean, cassava

and peanut.

22. Use of a transgenic plant according to any of claims 20 to 21 for the production of substances resulting from the expression of chimeric genes.

23. Use of a transgenic plant according to claim 22 wherein the substances are proteins, bioactive substances and oils.

24. Substances obtained according to any of claims 23 and 24.

TRATADO DE COOPERACIÓN EN MATERIA DE PATENTES

PCT

INFORME DE BÚSQUEDA INTERNACIONAL

(Artículo 18 y reglas 43 y 44 del PCT)

Referencia del expediente del solicitante o del mandatario	**PARA CONTINUAR LA TRAMITACIÓN** ver la notificación de transmisión del informe de búsqueda internacional (Formulario PCT/ISA/220) y, en su caso, el punto 5 de esta hoja.	
Solicitud internacional nº PCT/ES 99/00017	Fecha de presentación internacional *(día/mes/año)* 23 Enero 1999 (23.01.99)	Fecha de prioridad (la más antigua) *(día/mes/año)* 23 Enero 1998 (23.01.98)
Solicitante	CONSEJO SUPERIOR INVESTIGACIONES CIENTIFICAS, PRIETO DAPENA, Pilar y otros	

El presente informe de búsqueda internacional, elaborado por esta Administración encargada de la Búsqueda Internacional, se transmite al solicitante, conforme al artículo 18. Se remite una copia del mismo a la Oficina Internacional.

Este informe de búsqueda internacional comprende un total de 3 hojas.

☐ Se adjunta una copia de cada uno de los documentos citados en el informe relativos al estado de la técnica.

1. **Consideraciones sobre el informe**
   a. En lo que se refiere al **idioma**, la búsqueda internacional se ha realizado sobre la solicitud internacional en el idioma en el cual se depositó, salvo indicación en contra señalada en este apartado.

   ☐ la búsqueda internacional se ha realizado sobre una traducción de la solicitud internacional facilitada a esta Administración (Regla 23.1 b)).

   b. En lo que se refiere a **las secuencias de nucleótidos y/o de aminoácidos** divulgadas en la solicitud internacional (en su caso), la búsqueda internacional se ha basado en la lista de secuencias:

   ☐ contenida en la solicitud internacional en formato escrito.

   ☐ presentada conjuntamente con la solicitud internacional en soporte legible por ordenador.

   ☒ facilitada posteriormente a esta Administración por escrito.

   ☐ facilitada posteriormente a esta Administración en soporte legible por ordenador.

   ☒ se ha entregado la declaración, según la cual la lista de secuencias presentada por escrito posteriormente no va más allá de la divulgación hecha en la solicitud internacional tal y como fue presentada.

   ☐ se ha entregado la declaración, según la cual la información grabada en el soporte legible por ordenador es idéntica a la lista de secuencias presentada por escrito.

2. ☐ **Se estima que algunas reivindicaciones no pueden ser objeto de búsqueda** (ver recuadro I).

3. ☐ **Falta unidad de invención** (ver recuadro II).

4. Con respecto al **título**,

   ☒ el texto se aprueba según fue remitido por el solicitante.

   ☐ el texto ha sido establecido por esta Administración con la siguiente redacción:

5. Con respecto al **resumen** ,

   ☒ el texto se aprueba según fue remitido por el solicitante.

   ☐ el texto (reproducido en el recuadro III) ha sido establecido por esta Administración de conformidad con la regla 38.2b). El solicitante puede presentar observaciones a esta Administración en el plazo de un mes a contar desde la fecha de expedición del presente informe de búsqueda internacional.

6. La figura de los **dibujos** a publicar junto con el resumen es la siguiente: Figura nº _____

   ☐ propuesta por el solicitante. ☒ N debe publicarse ninguna figura.

   ☐ por no haber propuesto el solicitante ninguna figura.

   ☐ por caracterizar mej r, esta figura, la invención.

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional nº PCT/ ES 99/00017

A. CLASIFICACIÓN DEL OBJETO DE LA SOLICITUD

CIP[6] C12N 15/82, C12N 15/29, A01H 5/00
De acuerdo con la Clasificación Internacional de Patentes (CIP) o según la clasificación nacional y la CIP.

B. SECTORES COMPRENDIDOS POR LA BÚSQUEDA

Documentación mínima consultada (sistema de clasificación, seguido de los símbolos de clasificación)

CIP[6] C12N, A01H

Otra documentación consultada, además de la documentación mínima, en la medida en que tales documentos formen parte de los sectores comprendidos por la búsqueda

Bases de datos electrónicas consultadas durante la búsqueda internacional (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

CAS, WPI, EPODOC

C. DOCUMENTOS CONSIDERADOS RELEVANTES

Categoría*	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones nº
X	ALMOGUERA et al. "Developmental and environmental concurrent expression of sunflower dry-seed-stored low-molecular-weight heat-shock protein and Lea mRNAs". 1992 Plant Mol. Biol. Vol. 19(5). Págs. 781-92	1-3
A	WO 9713843 A (CORNELL RESEARCH FOUNDATION INC.) 17.04.1997, pág. 3, línea 13 - pág. 7, línea 25	1-13
A	HULL, G. et al "Analysis of the promoter of an abscisic acid responsive late embryogenesis abundant gene of Arabidopsis thaliana" 1996. Plant Sci. Vol. 114(2). Págs. 181-92	1-13

☐ En la continuación del recuadro C se relacionan otros documentos ☒ Los documentos de familia de patentes se indican en el anexo

* Categorías especiales de documentos citados:

"A" documento que define el estado general de la técnica no considerado como particularmente relevante.

"E" solicitud de patente o patente anterior pero publicada en la fecha de presentación internacional o en fecha posterior.

"L" documento que puede plantear dudas sobre una reivindicación de prioridad o que se cita para determinar la fecha de publicación de otra cita o por una razón especial (como la indicada).

"O" documento que se refiere a una divulgación oral, a una utilización, a una exposición o a cualquier otro medio.

"P" documento publicado antes de la fecha de presentación internacional pero con posterioridad a la fecha de prioridad reivindicada.

"T" documento ulterior publicado con posterioridad a la fecha de presentación internacional o de prioridad que no pertenece al estado de la técnica pertinente pero que se cita por permitir la comprensión del principio o teoría que constituye la base de la invención.

"X" documento particularmente relevante; la invención reivindicada no puede considerarse nueva o que implique una actividad inventiva por referencia al documento aisladamente considerado.

"Y" documento particularmente relevante; la invención reivindicada no puede considerarse que implique una actividad inventiva cuando el documento se asocia a otro u otros documentos de la misma naturaleza, cuya combinación resulta evidente para un experto en la materia.

"&" documento que forma parte de la misma familia de patentes.

Fecha en que se ha concluido efectivamente la búsqueda internacional. 19 Mayo 1999 (19.05.1999)	Fecha de expedición del informe de búsqueda internacional 2 JUN 1999 (02.06.99)
Nombre y dirección postal de la Administración encargada de la búsqueda internacional O.E.P.M. C/Panamá 1, 28071 Madrid, España. nº de fax +34 91 3495304	Funcionario autorizado A. Collados Martín- Posadillo nº de teléfono + 34 91 3495552

Formulario PCT/ISA/210 (segunda hoja) (julio 1998)

INFORME DE BÚSQUEDA INTERNACIONAL
Información relativa a miembros de familias de patentes

Solicitud internacional nº
PCT/ ES 99/00017

Documento de patente citado en el informe de búsqueda	Fecha de publicación	Miembro(s) de la familia de patentes	Fecha de publicación
WO 9713843 A	17.04.1997	CA 2234168 A	17.04.1997
		EP 0874897 A	04.11.1998
		AU 7397796 A	30.04.1997

PCT ORGANIZACION MUNDIAL DE LA PROPIEDAD INTELECTUAL
Oficina Internacional

SOLICITUD INTERNACIONAL PUBLICADA EN VIRTUD DEL TRATADO DE COOPERACION EN MATERIA DE PATENTES (PCT)

(51) Clasificación Internacional de Patentes 6 :
C12N 15/82, 15/29, A01H 5/00

A1

(11) Número de publicación internacional: **WO 99/37795**

(43) Fecha de publicación internacional: 29 de Julio de 1999 (29.07.99)

(21) Solicitud internacional: PCT/ES99/00017

(22) Fecha de la presentación internacional: 23 de Enero de 1999 (23.01.99)

(30) Datos relativos a la prioridad:
P 9800122 23 de Enero de 1998 (23.01.98) ES

(71) Solicitante *(para todos los Estados designados salvo US)*: CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS [ES/ES]; Calle Serrano, 117, E–28006 Madrid (ES).

(72) Inventores; e
(75) Inventores/solicitantes *(sólo US)*: PRIETO–DAPENA, Maria Pilar [ES/ES]; Instituto Recursos Naturales y Agrobiología Sevilla, Consejo Superior de Investigaciones Científicas, Apartado 1052 Estafeta–Puerto, E–41080 Sevilla (ES). ALMOGUERA ANTOLINEZ, María Concepción [ES/ES]; Instituto Recursos Naturales y Agrobiología Sevilla, Consejo Superior de Investigaciones Científicas, Apartado 1052 Estafeta–Puerto, E–41080 Sevilla (ES). JORDANO FRAGA, Juan Bautista [ES/ES]; Instituto Recursos Naturales y Acrobiología Sevill, a, Consejo Superior de Investigaciones Científicas, Apartado 1052 Estafeta–Puerto, E–41080 Sevilla (ES).

(74) Mandatario: OJEDA GARCIA, Pedro; Consejo Superior de Investigaciones Científicas, Calle Serrano, 113, E–28006 Madrid (ES).

(81) Estados designados: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, Patente ARIPO (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Patente euroasiática (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Patente europea (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), Patente OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Publicada
Con informe de búsqueda internacional.
Antes de la expiración del plazo previsto para la modificación de las reivindicaciones, será publicada nuevamente si se reciben modificaciones.

(54) Title: PROMOTER AND REGULATOR SEQUENCES *Ha ds10 G1*: A GENE LEA OF SUNFLOWER EXPRESSED EXCLUSIVELY IN SEEDS FROM THE MATURATION PHASE

(54) Título: PROMOTOR Y SECUENCIAS REGULADORAS DE *Ha ds10 G1*: UN GEN LEA DE GIRASOL EXPRESADO EXCLUSIVAMENTE EN SEMILLAS DESDE LA FASE DE MADURACION

(57) Abstract

The present invention discloses the isolation and characterization in transgenic tobacco plants of the promoter and regulator sequences of a gene LEA–I of sunflower, *Ha ds10 G1*. These sequences present characteristics which are extremely appropriate to be used in the modification of seeds (for example of reserve substances). The advantages of their possible use in transgenic plants are demonstrated through examples such as studies related to the accumulation and location of RNAm Ha ds10 in the homologous system. Said studies show both the high expression levels reached during embryogenesis from the early maturation phases and the absolute specificity of the seed, together with a homogenous location in embryos which is finally restricted essentially to the soft tissue in palisade of the cotyledons, a tissue specialized in the accumulation of reserve substances in the sunflower.

(57) Resumen

Con la presente invención aislamos y caracterizamos en plantas transgénicas de tabaco, el promotor y las secuencias reguladoras de un gen LEA–I de girasol, *Ha ds10 G1*. Estas secuencias presentan unas características muy apropiadas para su uso en la modificación de semillas (por ej. de sustancias de reservas). Las ventajas de su posible uso en plantas transgénicas se muestran mediante ejemplos como estudios de la acumulación y localización del ARNm Ha ds10 en el sistema homólogo. Estos estudios muestran tanto los elevados niveles de expresión alcanzados durante la embriogénesis desde fases tempranas de la maduración, como sus absoluta especificidad de semilla, acompañada de una localización homogénea en embriones que acaba restringiéndose fundamentalmente al parénquima en empalizada de los cotiledones, un tejido especializado en la acumulación de sustancias de reservas en el girasol.

UNICAMENTE PARA INFORMACION

Códigos utilizados para identificar a los Estados parte en el PCT en las páginas de portada de los folletos en los cuales se publican las solicitudes internacionales en el marco del PCT.

AL	Albania	**ES**	España	**LS**	Lesotho	**SI**	Eslovenia
AM	Armenia	**FI**	Finlandia	**LT**	Lituania	**SK**	Eslovaquia
AT	Austria	**FR**	Francia	**LU**	Luxemburgo	**SN**	Senegal
AU	Australia	**GA**	Gabón	**LV**	Letonia	**SZ**	Swazilandia
AZ	Azerbaiyán	**GB**	Reino Unido	**MC**	Mónaco	**TD**	Chad
BA	Bosnia y Herzegovina	**GE**	Georgia	**MD**	República de Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tayikistán
BE	Bélgica	**GN**	Guinea	**MK**	Ex República Yugoslava de Macedonia	**TM**	Turkmenistán
BF	Burkina Faso	**GR**	Grecia			**TR**	Turquía
BG	Bulgaria	**HU**	Hungría	**ML**	Malí	**TT**	Trinidad y Tabago
BJ	Benin	**IE**	Irlanda	**MN**	Mongolia	**UA**	Ucrania
BR	Brasil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarús	**IS**	Islandia	**MW**	Malawi	**US**	Estados Unidos de América
CA	Canadá	**IT**	Italia	**MX**	México	**UZ**	Uzbekistán
CF	República Centroafricana	**JP**	Japón	**NE**	Níger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Países Bajos	**YU**	Yugoslavia
CH	Suiza	**KG**	Kirguistán	**NO**	Noruega	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	República Popular Democrática de Corea	**NZ**	Nueva Zelandia		
CM	Camerún			**PL**	Polonia		
CN	China	**KR**	República de Corea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstán	**RO**	Rumania		
CZ	República Checa	**LC**	Santa Lucía	**RU**	Federación de Rusia		
DE	Alemania	**LI**	Liechtenstein	**SD**	Sudán		
DK	Dinamarca	**LK**	Sri Lanka	**SE**	Suecia		
EE	Estonia	**LR**	Liberia	**SG**	Singapur		

TÍTULO

PROMOTOR Y SECUENCIAS REGULADORAS DE *HA DS10 G1:* UN GEN LEA DE GIRASOL EXPRESADO EXCLUSIVAMENTE EN SEMILLAS DESDE LA FASE DE MADURACIÓN.

SECTOR DE LA TÉCNICA

Agricultura. Esta invención se relaciona con la obtención de secuencias de ADN reguladoras ("promotores") y la construcción, usando dichas secuencias, de nuevos genes quiméricos capaces de expresarse de forma específica en semillas de plantas transgénicas. El gen *Ha ds10 G1* tiene la peculiaridad de expresarse exclusivamente en semillas de girasol desde la fase de maduración hasta la de desecación; sin responder a hormonas como el ácido abscísico (ABA), o al estrés hídrico en tejidos vegetativos. Además, el gen *Ha ds10 G1* se expresa de forma homogénea en embriones inmaduros, y preferentemente en el parénquima en empalizada de los cotiledones de embriones maduros. Estos patrones de expresión, junto con los elevados niveles de actividad del gen, sugieren que sus secuencias reguladoras sean especialmente adecuadas para la manipulación genética de sustancias de reserva en semillas

ESTADO DE LA TÉCNICA

Para conferir expresión específica en semillas de plantas transgénicas, hasta el momento se han aislado, caracterizado y utilizado promotores pertenecientes sobre todo a genes vegetales que codifican proteínas de reserva, u otros productos expresados exclusivamente en semillas durante diversas etapas del desarrollo [véanse por ejemplo las siguientes referencias bibliográficas y patentes, así como otros documentos citados en ellas: Thomas TL, en *Plant Cell*, vol 5, pp 1401-1410, 1993; Gatehouse JA, y Shirsat AH, en *Control of Plant Gene Expression*, pp 357-375, *CRC press*, 1993; y las patentes USA números: 5530192, 5530194 y 5420034]. Esto ha permitido por ejemplo la obtención de nuevas plantas transgénicas con semillas modificadas en su contenido de ácidos grasos y de prot ínas de reserva [veáns por ejemplo:

Voelker TA, Worrell AC, Anderson L,Bleibaum J, Fan C, Hawkins DJ, Radke SE y Davies HM, en *Science*, vol. 257, pp.72-74, 1992; y Saalbach I, Pickardt T, Machemehl F, Saalbach G, Schieder O, y Muntz K, en *Molecular and General Genetics* 242: 226-236, 1994]. Para el desarrollo del enorme potencial de esta técnica, pudieran ser útiles otros promotores con distintas especifidades de tejido en la semilla y diversos patrones temporales de expresión. Recientemente en nuestro grupo, y otros laboratorios, hemos descrito la expresión en semillas de genes que codifican proteínas de choque térmico de bajo peso molecular (sHSPs: *small heat-shock proteins*). Uno de estos genes, *Ha hsp17.7 G4* , muestra, en plantas transgénicas de tabaco, patrones de expresión adecuados para su posible uso en la modificación de semillas mediante ingeniería genética: dicho gen se expresa desde etapas tempranas de la maduración de la semilla, y con una especificidad de tejido asociada a los cotiledones [Coca MA, Almoguera C, Thomas TL, y Jordano J, en: *Plant Molecular Biology* 31: 863-876, 1996]. Sin embargo el gen *Ha hsp17.7 G4*, al igual que otros genes vegetales sHSP expresados en semillas, también se expresa en respuesta al calor (choques térmicos) en tejidos vegetativos de la planta tras la germinación de las semilla. Esto último imposibilita su uso en ingeniería genética cuando se requieren secuencias de ADN reguladoras que garanticen que no haya expresión de los genes quiméricos fuera de la semilla: por ejemplo, cuando la expresión fuera de lugar de estos genes pueda afectar a la viabilidad, el crecimiento o la salubridad de las plantas transgénicas. Para solucionar estos problemas hemos modificado las secuencias reguladoras del gen *Ha hsp17.7 G4* de forma que genes quiméricos que contengan estas secuencias mantengan su expresión en semillas y pierdan su inducción por calor; procedimiento utilizable para la modificación y uso similar de secuencias reguladoras de otros genes sHSP expresados en semillas [Almoguera, Prieto-Dapena y Jordano, solicitud de patente #9602746 (Oficina Española de Patentes)]. De forma alternativa, también hemos propuesto un uso similar para el promotor y las secuencias reguladoras del gen de girasol *Ha hsp17.6 G1*, que únicamente se expresa en semillas. Dicho gen no responde al calor o a otro tipo de estrés (frío, desecación, tratamiento hormonal con ABA) en tejidos vegetativos [Carranco,

Almoguera y Jordano, solicitud de patente #9701215 (Oficina Española de Patentes).

En la presente solicitud proponemos usos análogos alternativos para el promotor y las secuencias reguladoras del gen LEA de girasol *Ha ds10 G1*. El gen *Ha ds10 G1* está incluido en un clon genómico correspondiente a un ADNc descrito previamente (*Ha ds10*, número de acceso X50699) cuyos patrones de expresión se conocían de forma incompleta [Almoguera y Jordano, *Plant Mol. Biol.* 19:781-792, 1992]. El promotor y secuencias reguladoras de este gen (*Ha ds10 G1*) han sido clonados y se describen, caracterizan y utilizan por primera vez en los ejemplos de esta solicitud. El gen *Ha ds10 G1* pertenece a la familia de genes LEA (*Late Embryogenesis Abundant*) de Clase I (tipo D-19 ó LEA-I). Estos genes codifican proteínas altamente conservadas en varias especies vegetales, y su expresión está generalmente restringida a semillas y a fases tempranas de la germinación [ver por ejemplo las siguientes revisiones: Dure III, L., *Structural motifs in Lea proteins*, en *Plant Responses to Plant Dehydration During Environmental Stress.*, Close TJ and Bray EA Eds., *Current Topics in Plant Physiology 10*: 91-103, 1993; y Delseny M, Gaubier P, Hull G, Saez-Vasquez J, Gallois P, Raynal M, Cooke R, Grellet F., *Nuclear Genes expressed during seed desiccation: relationship with responses to stress* , en Stress-induced Gene Expression in Plants (Basra, A. S., ed.), pp. 25-59, Harwood Academic Publishers, Reading, 1994]. Los promotores de los genes LEA no han sido considerados como buenos candidatos para su uso en proyectos de modificación de sustancias de reserva en semilla, ya que en general presentan actividad en fases posteriores a la maduración de la semilla, como durante la desecación del embrión [ver las consideraciones de Kridl JC, Knauf VC, Thompson GA, en *Control of Plant Gene Expression.* pp. 481-498, CRC press,1993]. Sin embargo se conocen genes LEA que se activan en fases de maduración anteriores a la desecación, como los genes de algodón denominados LEA-A [Hughes DW y Galau GA, The Plant Cell 3:605-618, 1991]. También dentro los genes LEA de clase I se conocen ejemplos de activación anterior a la desecación, como en el caso de los genes *At Em1*, *emb564*, y *emb1* [respectivamente en arabidopsis, maíz y zanahoria: Gaubier P, Raynal M,

Hull G, Huestis GM, Grellet F, Arenas C, Pages M, y Delseny M, *Mol. Gen. Genet.*, 238: 409-418, 1993; Williams B, y Tsang A, *Plant Mol. Biol.*, 16: 919-923, 1991; Wurtele ES, Wang H, Durgerian S, Nikolau BJ, y Ulrich TH. *Plant Physiol.* 102:303-312, 1993]. Estos ejemplos indicarían el posible uso de secuencias reguladoras de genes de esta familia para la modificación de semillas. No obstante, su uso concreto estaría limitado tanto por los niveles de expresión alcanzados en cada caso y en cada fase del desarrollo; como por las distintas especificidades de tejido. Así por ejemplo, aunque en *Arabidopsis* el gen *At Em1* se activa tempranamente, su expresión esta fundamentalmente restringida a tejidos provasculares de los cotiledones y a tejidos corticales externos del eje embrionario [Gaubier, P., Raynal, M., Hull, G., Huestis, GM., Grellet, F., Arenas, C., Pages, M., y Delseny, M., Mol. Gen. Genet., 238: 409-418, 1993]. En el caso del gen *emb1* de zanahoria, sus ARNm se localizan preferente en los meristemos del embrión, particularmente en el *procambium* [Wurtele ES, Wang H, Durgerian S, Nikolau BJ, y Ulrich TH. *Plant Physiol.* 102:303-312, 1993]. No se han publicado las secuenciasgenómicas del gen*emb564*, y se desconoce la localización precisa de sus ARNm [Williams B, y Tsang A, *Plant Mol. Biol.*, 16: 919-923, 1991].

La expresión del gen de girasol *Ha ds10 G1*, así como su promotor y secuencias reguladoras presentan, como se describe a continuación, unas características únicas entre las de otros miembros de la familia LEA-I; lo que hace que dichas secuencias sean potencialmente utilizables en la modificación de semillas mediante ingeniería genética.

DESCRIPCION DE LA INVENCION

Con la presente invención aislamos y caracterizamos en plantas transgénicas de tabaco, el promotor y las secuencias reguladoras de un gen LEA-I de girasol, *Ha ds10 G1*. Estas secuencias (Ejemplo 1) presentan unas características muy apropiadas para su uso en la modificación de semillas (por ej. de sustancias de reservas). Las ventajas de su posible uso en plantas transgénicas se muestran mediante otros ejemplos: A.- Estudios de la acumulación y localización del ARNm Ha ds10 en el sistema homólogo (Ejemplo

2). Estos estudios muestran tanto los el vados niveles d expresión alcanzados durante la embriogénesis desde fases tempranas de la maduración, como sus absoluta especificidad de semilla, acompañada de una localización homogénea en embriones que acaba restringiéndose fundamentalmente al parénquima en empalizada de los cotiledones, un tejido especializado en la acumulación de sustancias de reservas en el girasol. B.- En el ejemplo 3, ilustramos también el posible uso de dichas secuencias mediante la construcción y análisis en plantas transgénicas de distintos genes quiméricos; usando el promotor y combinaciones de distintas secuencias reguladoras de *Ha ds10 G1* (5'-flanqueantes, codificantes, intrón y 3'-flanqueantes), con el gen indicador (*reporter*) de la ß-glucuronidasa bacteriana (GUS). Estos ejemplos demuestran en un sistema heterólogo modelo (tabaco) la utilidad de los distintos genes quiméricos ensayados: alto nivel de expresión y especifidad de semillas desde fases tempranas de la maduración, así como la contribución funcional de las distintas secuencias ensayadas. Mediante los ejemplos adjuntos mostramos que la especifidad de semillas está conferida fundamentalmente por el promotor y secuencias *5'*-flanqueantes de *Ha ds10G1* (incluyendo secuencias no-transcritas y transcritas: como el 5'-UTR y parte de la secuencia codificante). Adicionalmente las secuencias 3'-flanqueantes incrementan los niveles de expresión en semillas; y el intrón los reduce de forma específica en tejidos no-embrionarios. Dada la conservación de la regulación de la expresión de genes embrionarios en semillas de plantas, incluidos los genes LEA-I [Thomas TL, en *The Plant Cell* 5:1401-1410, 1993]; estas secuencias podrían usarse tanto en el sistema homólogo (el girasol) como en otros sistemas heterólogos de gran importancia económica (por ejemplo la colza, la soja, el maíz, etc).

La realización práctica de esta invención, representada con los ejemplos y figuras adjuntos, utiliza técnicas convencionales de Biología Molecular, Microbiología, ADN recombinante; y de producción de plantas transgénicas, que son de uso común en laboratorios especializados en estos campos. Estas técnicas están explicadas con suficient detalle en la literatura científica [veánse por ejemplo: Sambrok J, Fritsch EF, y Maniatis T, *Molecular Cloning: A Laboratory Manual, Cold Spring Harbor laboratory Press*, 2ª Edición, 1989;

Glover DM, *DNA Cloning, IRL Press* , 1985; Lindsey K.,*Plant Tissue Culture Manual, Kluwer Academic Publishers*, 1993; y Gelvin SB, Schilperoort RA, Verma DPS, Plant Molecular Biology Manual, *Kluwer Academic Publishers* , 1992]. Para otros detalles mas específicos, se citan las referencias bibliográficas pertinentes en el lugar correspondiente de esta solicitud.

EJEMPLO 1: clonación, determinación del mapa de restricción, secuencia nucleotídica, y análisis del promotor de *Ha ds10 G1*.

Para obtener el clon *Ha ds10 G1* se rastreó la genoteca de ADN genómico de girasol descrita por Coca *et al.* [Plant Mol. Biol. 31: 863-876, 1996], con la sonda correspondiente al ADNc completo *Ha ds10* [Almoguera y Jordano, Plant Mol. Biol. 19: 781-792, 1992]; usando las condiciones de hibridación y procedimientos estandard de clonación molecular descritos con suficiente detalle en la primera de estas referencias (Coca *et al.*, 1996). Así, aislamos un fago (IGEM11) con un inserto de ADN genómico de girasol de aproximadamente 16.5 Kb cuyo mapa parcial se muestra en la Figura 1. Mediante análisis de restricción, determinamos que dos fragmentos adyacentes de Sac I (de 4.2 y 9.3 Kb) contienen las secuencias que hibridan con el ADNc. Se determinó un mapa de restricción detallado del primero de estos fragmentos, y de parte (Å4 Kb) del segundo (Figura 1). Distintos subfragmentos de ADN genómico, correspondientes a la región mapeada, se clonaron en el vector pBluescript SK+, dando lugar a los plásmidos cuyo nombre e inserto se indica en la Figura 1. A partir de estos plásmidos se determinó, en ambas cadenas del ADN y por el método de Sanger (*dideoxi*), la secuencia nucleotídica de 3617 bp entre los sitios de Sac I y Sma I (Figura 1, parte inferior). Estos datos se presentan en la SEQ Nº 1. Mediante comparaciones de secuencia confirmamos que parte de la secuencia genómica determinada se corresponde con la del ADNc Ha ds10 [Almoguera y Jordano, *Plant Mol. Biol.* 19: 781-792, 1992; número de acceso en GenBank X59699]. La secuencia de aminoácidos de la proteína codificada por el gen *Ha ds10 G1* se indica bajo las s cuencias nucleotídicas correspondient s. En el ADN genómico, la zona codificante está interrumpida por un intrón anómalamente largo (de 1024 bp), aunque situado en

una posición conservada n otros genes LEA d clase I [ver datos r visados por Simpson GC, Leader DJ, Brown JWS y Franklin T, en *Charasteristics of Plant pre-mRNA Introns and Transposable Elements*; *Plant Mol. Biol. LabFax*, pp. 183-252; Croy RRD Ed., Bios Scientific Publishers Ltd. 1993]. La única diferencia, entre las secuencias genómicas que codifican el ARNm y las del ADNc, fue una inversión de dos nucleótidos (GC en vez de CG) dentro del segundo exón (en las posiciones +1176 y +1177 desde el codón de iniciación); lo que provoca un cambio de un aminoacido (S en vez T) en la secuencia de la proteína. La diferencia se debe a un error (debido a una compresión) en la lectura inicial de las reacciones de secuencia del ADNc. Las secuencias de *Ha ds10 G1* que hemos determinado incluyen también 1576 bp, del promotor del gen y secuencias 5'-flanqueantes; y 553 bp de secuencias genómicas 3'-flanqueantes no presentes en el ADNc original.

Mediante la técnica de extensión del cebador (*primer extension*), se determinaron tres posibles sitios de iniciación de la transcripción en el promotor de *Ha ds10 G1*. Dos de estos sitios han sido confirmados mediante otras técnicas (sitios 1 y 2, indicados por flechas en la SEQ Nº 1). Para ello se utilizó, según el procedimiento descrito por Domon *et al.* [Domon C, Evrard JL, Pillay DTN, y Steinmetz A. *Mol. Gen. Genet.* 229:238-244, 1991], ARN total de embriones de girasol hibridado con el cebador sintético: 5'-CTCCTGTTCCGGAATTTTGCGTGT-3'; cuya secuencia corresponde a la de la cadena no codificante de *Ha ds10 G1*, entre las posiciones +25 y +48, desde el codón de iniciación. Las hibridaciones con el cebador se hicieron a 62ºC. Los híbridos se extendieron con transcriptasa reversa de AMV, durante 90 min. a 42ºC. Los productos de extensión se analizaron en geles de secuenciación PAGE al 6%, junto con reacciones de secuencia producidas usando el mismo cebador. Los sitios de iniciación 1 y 2 (en las posiciones -33 y -25; ver SEQ Nº 1) son funcionales, y se detectan de forma independiente usando la técnica de protección frente a la ribonucleasa A (RNAsa A, ver Figura 3A). Un tercer sitio de iniciación (sitio 3, en la posición -119 de la SEQ N 1) no fu confirmado claramente, mediante dicha técnica. Estos sitios de iniciación delimitan funcionalmente l extremo 3' del promotor del gen *Ha ds10 G1*.

El análisis de las secu ncias proximal s del promotor del gen *Ha ds10G1* mostró que dos de los sitios de iniciación detectados (los sitios 1 y 2) se encuentran a una distancia apropiada de una posible secuencia TATA (en la posición -86). El posible sitio mas distal (sitio 3, -119) no tiene secuencias TATA claras situadas en su proximidad. Además de estos elementos del promotor, se observaron dos posibles "cajas" RY (RY1 e RY2 en las posiciones -129 y -65 de la SEQ Nº 1), como las que participan en regulación de la expresión en semillas de numerosos genes de plantas [Dickinson CD, Evans RP, y Nielsen RC, en *Nucleic Acids Research* 16: 371, 1988].

Hemos modificado la caja RY1 situada en -129; verificando, mediante experimentos de expresión transitoria en embriones de girasol, su requerimiento funcional para la trans-activación del promotor de *Ha ds10G1* por factores transcripcionales de tipo ABI3 [Giraudat J., Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM en *The Plant Cell* 4: 1251-1261, 1992]. Para ello, preparamos modificaciones de las fusiones ds10::GUS construídas para estudios en plantas transgénicas (ver el Ejemplo 6.3 y la Figura 5). Los genes quiméricos contenidos en dos de estas fusiones (ds10F1y ds10F2) se purificaron como fragmentos de ADN que se subclonaron por ligación en el vector pBluescript SK+ (Promega); cambiando así las secuencias del vector binario por otras de menor tamaño, mas útiles para realizar experimentos de expresión transitoria. Así, usando el fragmento Sal I - Eco RI (con el gen quimérico obtenido a partir de ds10F1), obtuvimos el plásmido pSKds10F1. En el caso de ds10F2, el fragmento de Sph I - Eco RI (desde la posición -125 en *Ha ds10 G1*, hasta el extremo 3' de *nos*) se ligó al fragmento complementario (que contiene el promotor y secuencias 5´flanqueantes de *Ha ds10 G1*), purificado tras la digestión de pSKds10F1 con Sph I y Eco RI, resultando en el plásmido pSKds10F2. Finalmente a partir de los plásmidos pSKds10F1 y pSKds10F2 (mapas no mostrados) se obtuvieron versiones mutagenizadas de los mismos tras la digestión de su ADN con Sph I, haciendo romos los extremos resultantes mediant tratamientos con ADN polimerasa de T4, seguidos de re-ligación del ADN. De esta forma obtuvimos los plásmidos pSKds10F1ÆRY y pSKds10F2ÆRY (mapas no mostrados). Estos plásmidos difieren únicamente en una deleción de 5 nucleótidos entre las

posiciones -126 y -122 del promotor de *Ha ds10 G1*. Estos cambios destruyeron la caja RY1 presente en los genes quim´ricos ds10F1 y ds10F2 (ver Figuras 1, 2 y 5), lo que se verificó mediante reacciones de secuenciación por el método de Sanger (dideoxy), usando el cebador 5'CTCCTGTTCCGGAATTTTGCGTGT3' (cadena no codificante de *Ha ds10G1*, entre las posiciones +25 y +48).

Los experimentos de trans-activación en expresión transitoria se realizaron mediante el bombardeo de embriones de girasol con proyectiles cubiertos de mezclas de ADN de distintos plásmidos. Estas mezclas contienen un plásmido de referencia, pDO432 [Ow DW, Wood KV, deLuca M, de Wet JR, Helinski D y Howell SH. Science 234: 856-859, 1996], con el gen de la luciferasa (LUC) de luciérnaga (*Photinus pyralis*) bajo el control del promotor CaMV 35S; la fusión ds10::GUS ensayada en cada caso (con las secuencias RY1 intactas o modificadas),y un plásmido efector, pABI3, que expresa el factor ABI3 bajo el control del promotor CaMV 35S. pABI3 se obtuvo sustituyendo el ADNc de Pv ALF en el plásmido pALF [Bobb AJ, Eiben HG, y Bustos MM en *The Plant Journal* 8: 331-343, 1995], por el ADNc de ABI-3. El ADNc de ABI3 se clonó como un fragmento Xba I (hecho romo con *klenow*) - Eco RI (parcial), purificado a partir del plásmido pcabi3-4F [Giraudat J., Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM en *The Plant Cell* 4: 1251-1261, 1992]. El plásmido pABI3 se añade a la mezcla, o se omite, para probar el efecto del factor ABI3 sobre la expresión GUS de la fusión ensayada. Los experimentos se realizaron esencialmente como se describe por Bobb *et al.*, [Bobb AJ, Eiben HG, y Bustos MM en *The Plant Journal* 8: 331-343, 1995], con las siguientes modificaciones. Los embriones de girasol (17-20 dpa) se prepararon como sigue. Las semillas de girasol se esterilizan con lavados en etanol 70% durante 1 min, y en 2% de hipoclorito sódico con una gota de Tritón X-100 durante 40 min, finalizados con varios lavados con agua destilada; tras los que se pelan en condiciones estériles. Los embriones se cortan longitudinalmente (separando sus dos cotiledones) y se colocan con la superficie cortada, sobre placas con medio sólido MS, que contiene 2% sacarosa y 0.5 M sorbitol. A continuación se precultivan durant 2-4 h en oscuridad y temperatura ambi nte (25°C). Todos los plásmidos fueron purificados usando el *Quantum midiprep kit* (Biorad).

Normalmente se usaron para cada disparo: 0.2 µg del plásmido de referencia, 1 µg del plásmido ds10::GUS y 1 µg del plásmido efector (o la misma cantidad del plásmido pJIT82 en los controles negativos). Para la preparación de las partículas de oro, así como la precipitación del ADN sobre las mismas, se siguió el método descrito por Chern *et al.* [Chern MS, Bobb AJ y Bustos M. *The Plant Cell* 8: 305-321, 1996]. El bombardeo de partículas se llevó a cabo con el sistema *Biolistic PDS-1000 He* (Biorad). Las condiciones de bombardeo fueron: Membrana de ruptura de 1550 psi, partículas de oro de 1.6 µm de diámetro, distancia de la membrana de ruptura al *macrocarrier* de 8 mm, distancia del *macrocarrier* a la rejilla de 6 mm; y distancia al tejido a bombardear de 6 cm. Los cotiledones bombardeados se incubaron durante 24 h a 28 °C en la oscuridad; tras lo cual se ensayó la actividad GUS (referida a la actividad LUC), como se describe por Bobb *et al.* [Bobb AJ, Eiben HG, y Bustos MM en *The Plant Journal* 8: 331-343, 1995].

La adición del plásmido efector pABI3 tuvo un efecto claro sobre la expresión relativa de GUS/LUC en bombardeos con la fusión pSKds10F2 (incremento medio de actividad relativa Å46.2X). En cambio, si la trans-activación se hace con el mismo plásmido mutado en la caja RY1 (pSKds10F2ÆRY1), se observó un descenso significativo del incremento medio de actividad relativa debido al efecto de ABI3 (Å26.3 X). Este resultado, mostrado en la figura 2, confirma el requerimiento funcional de la secuencia RY1 (posición -129 en la SEQ Nº 1). Por lo tanto esta caja RY participa en la activación transcripcional en semillas del promotor *Ha ds10 G1*, por factores del tipo ABI3 [Giraudat J., Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM en *The Plant Cell* 4: 1251-1261, 1992]. Otras secuencias del promotor (por ej. RY2 en -65) tambien pudieran contribuir al efecto de transactivación observado, ya que la mutación ensayada no destruye completamente el efecto activador de ABI3.

EJEMPLO 2: Acumulación y localización específica del mRNA *Ha ds10* en embriones de girasol:

Los patrones de acumulación d los ARN mensajeros del gen *Ha ds10G1* se determinaron mediante la t´cnica de la protección frente a la Ribonucleasa A (RNAsa A), descrita con detalle por Almoguera *et al.* [Almoguera C, Coca MA, Jordano J. *Plant Physiol.* 107: 765-773, 1995].Para ello, se utilizaron muestras de ARN total preparadas a partir de embriones de semillas en distintos estados de desarrollo en condiciones normales de crecimiento [Almoguera y Jordano, Plant Mol. Biol. 19: 781-792, 1992; Coca *et al.*, Plant Mol. Biol. 25: 479-492, 1994]; de germínulas de 3 días tras la imbibición (dpi); y de distintos órganos de plantas adultas antes de la floración. Los ARN de germínulas y plantas se prepararon a partir de material vegetal obtenido tanto en condiciones de crecimientro controlado [Almoguera y Jordano, *Plant Mol. Biol.* 19: 781-792, 1992; Coca MA, Almoguera C, y Jordano J. *Plant Mol. Biol.* 25: 479-492, 1994; Coca MA, Almoguera C, Thomas TL, y Jordano J. *Plant Mol. Biol.* 31: 863-876, 1996], como tras tratamientos de estrés: déficit de agua [Almoguera C, Coca MA, y Jordano J. *Plant J.* 4: 947-958, 1993; Coca MA, Almoguera C, Thomas TL, y Jordano J. *Plant Mol. Biol.* 31:863-876, 1996]; o tras la adición de hormonas como el ácido abscísico [Almoguera C y Jordano J. *Plant Mol. Biol.* 19: 781-792, 1992; Coca MA, Almoguera C, Thomas TL, y Jordano J. *Plant Mol. Biol.* 31: 863-876, 1996]. Las condiciones empleadas en cada tratamiento se describen con detalle en las referencias citadas en cada caso. La ribosonda usada para detectar los ARNm de *Ha ds10 G1* tiene una longitud de 396 nucleótidos, de los cuáles 63 son secuencias del vector pBluescript SK+ y el resto la secuencia de la cadena no-codificante de *Ha ds10 G1* entre las posiciones +212 y -121 (Sph I). Esta sonda hibrida con el extremo 5' de los ARN mensajeros de *Ha ds10 G1*, sobrepasando el sitio mas distal de iniciación de la transcripción (sitio 3, SEQ Nº 1), lo que permite detectar ARN mensajeros (ARNm) producidos a partir de los tres sitios de iniciación y la verificación experimental de las posiciones de iniciación. Esta ribosonda se preparó por transcripción *in vitro*, usando la ARN polimerasa T3, y como molde ADN del plásmido ds10G1S3Æ4.4. (Figura 1) que contiene las secu ncias de *Ha ds10G1* entre -1576 (Sal I) y +212, clonadas n el vector pBluescript SK+.

Los resultados en la Figura 3 muestran que los ARN mensajeros de *Ha ds10 G1* se detectan únicamente en semillas. Los niveles mayores d acumulación se observan en torno a 18-20 dpa, detectándose la expresión del gen a partir de los 10 dpa y desapareciendo tras la germinación (Figura 3). Los tratamientos con ABA, o déficit de agua no indujeron la acumulación de los ARN mensajeros de *Ha ds10 G1* (datos mostrados para ABA en germínulas; Figura 3). Como control positivo en las muestras de ARN analizadas para los distintos tratamientos, realizamos hibridaciones (datos no mostrados) con otra ribosonda de 651nucleótidos del gen *Ha hsp17.7 G4*, descrita anteriormente [Coca *et al.*, Plant Mol. Biol. 31: 863-876, 1996]; ya que dicho gen se expresa en respuesta a los distintos tratamientos ensayados. Estos análisis demostraron que los ARNm de *Ha ds10 G1* se acumulan exclusivamente en semillas, en condiciones normales del desarrollo y desde etapas tempranas de la maduración, confirmándose la iniciación a partir de al menos los sitios 1 y 2 (indicados en SEQ Nº 1). La banda indicada por el número 3 (Figura 3) no coincide bien con el tamaño esperado para el sitio de iniciación 3 (SEQ Nº1). Esta banda pudiera deberse a la protección de secuencias de ARN mensajeros de un gen muy homólogo; o bién del mismo *Ha ds10 G1*, conteniendo secuencias del intrón (ARNm sin procesar).

La distribución de los ARNm de *Ha ds10 G1* en embriones de girasol, fue investigada mediante experimentos de localización por hibridación *in situ*. Para ello los embriones se incluyeron en parafina, fijaron, seccionaron, e hibridaron con sondas específicas; esencialmente como se describe por Molinier [en la tesis: *Diplome D' Etudes Approfondies de Biologie Cellulaire et Moleculaire, Université Louis pasteur, Strasbourg*, 1995]. El tiempo de fijación se incrementó, desde 16 h a 4°C hasta 5 dias, aumentando según la edad de los embriones. La deshidratación de los embriones fijados se hizo por incubaciones sucesivas (2 veces cada una durante 30-90 min.) en etanol al 10%,20%,30%,40%,50%,60%,70%,95%, y 100%; seguidas de immersión en tolueno al 100% (1-3h, 2 veces). Los embriones fijados se incluyeron prim ro en tolu no:parafina (1:1), a 65°C durante 6-15 h, seguido de 5 inclusion s consecutivas en parafina, a 60°C durante 5-15 h. Las prehibridaciones e

hibridaciones con las sondas se hici ron a 45°C. La ribosonda específica de *Ha ds10 G1*, correspondiente al extremo 3'- del ARNm, se preparó como sigue. El plásmido ds10G1S1 (Figura 1) se usó como molde para preparar dos sondas por transcripción *in vitro* [Almoguera C, Coca MA y Jordano J. *Plant Physiol.* 107: 765-773, 1995] marcando con DIG-UTP. La sonda ds10-3'(-) se obtiene digiriendo el ADN del plásmido con Pvu II y efectuando la transcripción con ARN polimerasa T3. Esta sonda corresponde a la cadena no-codificante de *Ha ds10 G1* entre las posiciones +1202 (Pvu II en el segundo exón) y +1592 (extremo 3'). La segunda sonda [ds10-3' (+), usada como control], se preparó digiriendo el ADN de *Ha ds10 G1S1* con Bam HI (en el *polylinker*); y efectuando la transcripción con ARN polimerasa T7. La sonda ds10-3'(+) contiene la cadena codificante de *Ha ds10 G1*, entre las posiciones +870 y +1592. La especifidad de hibridación se determinó mediante experimentos de *Southern* similares a los descritos por Almoguera y Jordano [*Plant Mol. Biol.* 19: 781-792, 1992]. Mientras la hibridación con una sonda del ADNc completo detecta bandas correspondientes a unos 4-5 genes distintos en el genomio de girasol [Almoguera C, y Jordano J. *Plant Mol. Biol.* 19: 781-792, 1992]; usando la sonda ds10-3'(-) podemos detectar un único gen (con una ligera hibridación cruzada con otro; datos no mostrados).

Los resultados obtenidos en los experimentos de localización de ARN se muestran en la Figura 4. La sonda ds10-3'(-) es complementaria y de polaridad opuesta a los ARNm de *Ha ds10 G1*, lo que permite su detección. Los resultados obtenidos concuerdan con los datos de protección mostrados en la Figura 3, y muestran su acumulación en embriones desde los 12-15 dpa (Figura 4A) hasta los 21-28 dpa (Figuras 4C, F y H). Esta acumulación ocurre a niveles altos, lo que se deduce del corto tiempo preciso para su detección histoquímica (2-4 horas). En embriones inmaduros (Figura 4A) la distribución de los ARNm de Ha ds10 G1 es homogénea y comparable (Figura 4B) a la del ARNr 18S, que se detecta usando otra ribosonda correspondiente al fragmento G (Eco RI) del gen 18S de rábano [descrito por D Icasso-Tr mousaygue D, Grellet F, Panabieres F, Ananiev E D, y Dels ny, M. En Eur. J. Biochem. 172: 767-776, 1988]. En embriones mas maduros (21 dpa, Figura 4C) los ARNm de Ha ds10 G1 se

localizan también bastante homogéneamente, comenzando a detectarse una acumulación mas intensa en los haces vasculares (*procambium*), algo que no se observa con la sonda del ARNr 18S ni en éste ni en otros estadíos del desarrollo (Figuras 4D, B y G). Finalmente a los 28 dpa, los ARNm de *Ha ds10 G1* se localizan preferente en el parénquima en empalizada, un tejido especializado en la deposición de sustancias de reserva, situado en la cara interna de los cotiledones (Figuras 4F y H). Las localizaciones con la sonda ds10-3' (+), de la misma polaridad que los ARNm de *Ha ds10 G1*, no dieron señales de hibridación; lo que controló los experimentos descritos anteriormente (comparar las Figuras 4C y E). Estos experimentos demostraron que los patrones de expresión de los ARNm de *Ha ds10 G1* en girasol son muy especiales. La expresión observada en semillas, con altos niveles de acumulación desde etapas tempranas de la maduración embrionaria (10-12dpa), se combina con distribuciones espaciales que cambian desde la homogeinad hasta la mayor abundancia en tejidos de deposición de sustancias de reserva (parénquima en empalizada). La distribución y patrones de acumulación de los ARNm de *Ha ds10 G1* es distinta a la que presentan otros genes vegetales pertenecientes a la misma familia [Wurtele ES, Wang HQ, Durgerian S, Nikolau BJ y Ulrich TH. *Plant Physiol.* 102: 303-312, 1993; Gaubier, P., Raynal, M., Hull, G., Huestis, GM., Grellet, F., Arenas, C., Pages, M., y Delseny, M., *Mol. Gen. Genet.*, 238: 409-418, 1993]. Estos resultados indican la posible utilidad, para la modificación de semillas por Ingeniería genética, de genes quiméricos que incorporen las secuencias reguladoras de *Ha ds10 G1*.

<u>EJEMPLO 3: Construcción de genes quiméricos ds10G1::GUS y su análisis en plantas transgénicas de tabaco:.</u>

Como ejemplo para los posibles usos del promotor y las secuencias reguladoras del gen *Ha ds10 G1*,en la construcción de genes quiméricos con expresión específica en semillas de plantas transgénicas, describimos a continuación la construcción y el análisis en plantas transgénicas de tabaco de 4 fusiones traduccionales ds10G1::GUS (Figura 5). Dichas fusiones contienen, para su análisis funcional, el promotor y distintas combinaciones de secuencias

flanqueantes e intragénicas del gen *Ha ds10 G1*. Estas 4 fusiones proporcionan elevados niveles de expresión del gen indicador (GUS) en semillas desde etapas tempranas de la maduración (Figura 6), confirmando nuestras observaciones en el sistema homólogo (Ejemplo 2, Figuras 1-4).

La primera de estas construcciones, ds10F1 (Figura 5) se obtuvo a partir del plásmido ds10G1S3 (Figura 1), que contiene las secuencias genómicas de *Ha ds10 G1* entre Sal I (-1576) y Eco RI (+1086), subclonadas en los sitios de restricción correspondientes del vector pBluescript SK+ (Promega). Mediante tratamiento con Exonucleasa III del ADN de ds10G1S3 (previamente digerido con Hind III y Pst I), se delecionaron las secuencias de *Ha ds10 G1* entre Eco RI (+1086) y la posición +98 (en el primer exón), dando lugar al plásmido ds10G1S3Æ10.5 (Figura 1). Dicho plásmido de digirió con Bam HI (diana de restricción del *polylinker* situada inmediatamente adjacente a la posición +98 de *Ha ds10 G1*), rellenándose a continuación los extremos del ADN digerido usando el fragmento de Klenow de la ADN polimerasa I. A continuación el ADN se digirió con Sal I, purificándose el fragmento de 1679 p.b. que contiene las seuencias de *Ha ds10 G1* entre Sal I (-1576) y el extremo relleno de Bam HI. Este fragmento se clonó entre los sitios de Sal I y Sma I del vector binario pBI 101.2, resultando en ds10F1, una fusión traduccional que contiene 1576 nucleótidos de secuencias 5´-flanqueantes de *Ha ds10 G1* (desde el ATG) y los primeros 98 nucleótidos de la zona codificante, en fase con el gen GUS (Figura 5). La fusión ds10F2 se derivó a partir de ds10F1 mediante la inserción de un fragmento de ADN genómico de *Ha ds10G1* comprendido entre las posiciones (Figura 1) de +1205 (Pvu II), y Eco RI (Å+4670). Dicho fragmento contiene parte del segundo exón y Å3370 nucleótidos de secuencias 3'-flanquentes (a partir de codón de terminación en la posición +1301); y reemplaza a las secuencias nos-3' en la fusión ds10F1. El inserto Pvu II- Eco RI se purificó a partir de ADN del plásmido ds10G1S2 (Figura 1). Para la inserción de dicho fragmento, el ADN de ds10F1 se digirió con Sac I y los extremos del ADN se hicieron romos mediante tratamiento con la ADN polimerasa I de T4. A continuación, el ADN así tratado se digirió con Eco RI, purificándose el fragmento con las secuencias de *Ha ds10G1*. Este fragmento se ligó al inserto Pvu II- Eco RI anteriormente

descrito (con las secuencias 3'-flanqueantes de *Ha ds10 G1*), resultando en la fusión ds10F2 (Figura 4). La fusión ds10F2Æ (Figura 4) se obtuvo a partir de ds10F2, mediante la delección de las secuencias 3'-flanqueantes de *Ha ds10G1* entre Xba I (Å+2830) y Eco RI (Å+4670). Para ello, el ADN de ds10F2 se digirió con ambos enzimas; religándose, tras hacer romos los extremos de ADN resultantes con el fragmento de Klenow de la ADN polimerasa I. Finalmente, la cuarta fusión (ds10F3, Figura 5) se obtuvo a partir de un fragmento de ADN genómico de *Ha ds10 G1* entre Sal I (-1576) y Pvu II (+1204), purificado a partir del plásmido ds10G1S6 (Figura 1) tras la digestión con ambas enzimas de restricción. Este fragmento se ligó con ADN del vector pBI101.3, digerido previamente con Sal I y Sma I. La fusión ds10F3 contiene de esta forma el promotor y las mismas secuencias 5'-flanqueantes de *Ha ds10 G1* presentes en la fusión ds10F1, así como el primer exón (de +1 a +145), el intrón completo (de +146 a +1169) y parte del segundo exón de *Ha ds10 G1* (de+1170 a +1204), fusionado en fase con el gen GUS de pBI 101.3. En todos los casos la secuencia de nucleótidos correspondiente a la zona de fusión, entre las secuencias GUS y las de *Ha ds10 G1,* se comprobó mediante reacciones de secuenciación con el metodo de Sanger (dideoxi), usando como cebador las secuencias de GUS: 5'-ACGCGCTTTCCCACCAACGCTG-3'.

El ADN-T en las fusiones ds10F1, ds10F2, ds10F2Æ, y ds10F3 (Figura 5) fue movilizado desde *A. Tumefaciens* (LBA 4404), obteniéndose distintas plantas transgénicas de tabaco con integraciones independientes de cada gen quimérico. Estas plantas fueron obtenidas y caracterizadas mediante procedimientos estándard que se describen con detalle por Coca MA, Almoguera C, Thomas TL y Jordano J, [en *Plant Molecular Biology* , 31: 863-876, 1996]. En dichas plantas, la expresión del gen *GUS* se analizó tanto en semillas en desarrollo en condiciones normales de crecimiento (sin estrés exógeno); como en tejidos de germínulas, investigándose en este último caso los cambios de expresión inducidos por tratamientos con ABA y de desecación. Los análisis de semillas se realizaron con las plantas transgénicas originales (T0); mientras que para los de germínulas s utilizaron descendient s de estas plantas (T1), s gregantes para los genes quiméricos. S hicieron tanto estudios cuantitativos,

mediante el análisis fluorimétrico de los niveles de expresión *GUS* y de sus patrones temporales, como estudios cualitativos analizando histoquímicamente los patrones espaciales de expresión (especificidad de tejido). Estos estudios se hicieron como se describe con detalle por Coca MA, Almoguera C, Thomas TL y Jordano J, [en *Plant Molecular Biology*, 31: 863-876, 1996]. En total se obtuvieron y analizaron los siguientes números (entre paréntesis) de plantas transgénicas de tabaco, T0 "funcionales", con los genes quiméricos ds10F1 (14), ds10F2 (7), ds10F2Æ (8) y F3 (23). Estas plantas mostraron elevados niveles expresión del gen GUS en semilla (como consecuencia de la actividad del promotor y secuencias reguladoras del gen *Ha ds10 G1*), según se ilustra en la Figura 6 (paneles A-C). La integración de los distintos genes quiméricos en el ADN de las plantas transgénicas fue caracterizada mediante *Southerns* genómicos usando sondas de la región codificante de gen GUS; amplificaciones PCR de las secuencias próximas al empalme ds10::GUS, usando los cebadores 5'-ACGCGCTTTCCCACCAACGCTG-3' (GUS) y 5'-GAGTGAACAgAATtcCATCACAACAGGG-3' (ds10Eco RI); o mediante test de segregación de la resistencia a la Kanamicina (conferida por el gen *nptII*), relizados según se describe en [Jordano J, Almoguera C, y Thomas TL, *The Plant Cell* 1: 855-866, 1989]. Estos análisis determinaron que las plantas T0 seleccionadas para los estudios de expresión en semillas contenían de 1 a 5 integraciones independientes del gen quimérico correspondiente. La Figura 6 (adjuntada con esta solicitud) ilustra los resultados mas relevantes, obtenidos en el estudio de la expresión en plantas trangénicas de los genes quiméricos analizados. Estos resultados se describen con detalle a continuación.

La expresión GUS durante la maduración de las semillas en condiciones de crecimiento controladas (sin estrés exógeno), se analizó mediante ensayos fluorimétricos (Figura 6A) e histoquímicos (resumen en Figuras 6B-E). Los ensayos fluorimétricos se realizaron con semillas en estadíos definidos de maduración, a los 12, 16, 20, 24 y 28 días post-anthesis (dpa). Por cada planta T0 y estadío de maduración s preparon extractos de dos cápsulas florales distintas, y se ensayó la actividad GUS con Methilumbeliferilglucoronido (MUG) por duplicado (en total cuatro determinaciones de actividad por estadío de

desarrollo y planta transg´nica individual). La significación estadística de las diferencias observadas con las distintas fusiones GUS se determinó, tras la normalización logarítmica de los datos obtenidos, mediante análisis de la varianza [ANOVA, ver: Nap JP, Keizer P, y Jansen R, en *Plant Molecular Biology Reporter* 11: 156-164, 1993]. Los ensayos histoquímicos se hicieron con material diseccionado a partir de semillas, en estadíos de desarrollo definidos, procedentes de los siguiente números de plantas transgénicas: d10F1, 5; ds10F2, 6; ds10F2Æ, 6; y dsF3, 19. El endopermo y los embriones diseccionados a partir de semillas individuales se tiñeron con X-gluc, durante 150 min a 25ºC, analizándose de esta forma aproximadamente 150 semillas de cada planta transgénica.

Todos los genes quiméricos produjeron niveles elevados de expresión GUS en semillas, alcanzándose valores máximos medios de 1.65 x 10^6 pmol MU/ mg x min (Figura 6A: a los 24 dpa). Los ensayos histoquímicos confirmaron estos altos valores de actividad, ya que tanto los embriones (Figuras 6B y C) como el endospermo (Figura 6C) se tiñeron fuertemente a partir de los 12 dpa (Figura 6B), y con sólo 150 min de reacción. En ambos casos se observaron distribuciones espaciales de la actividad GUS bastante homogéneos (Figura 6B-C). Además, estos patrones de expresión no difirieron cualitativamente entre las plantas transgénicas de los distintos genes quiméricos (datos no mostrados).

Los ensayos fluorimétricos revelaron interesantes diferencias cuantitativas entre las distintas fusiones ds10::GUS. Estas diferencias dependen de las secuencias de *Ha ds10 G1* presentes en las fusiones. En algunos casos se ha podido mostrar la significación estadística de estas diferencias (con un nivel de confianza del 95%), lo que demuestra experimentalmente la contribución de las distintas secuencias ensayadas (promotor y secuencias 5'-flanqueantes, secuencias codificantes, 3'-flanqueantes, y del intrón) a los patrones de expresión embrionaria observados. La presencia en las fusiones de sequencias 3'-flanqueant s de *Ha ds10 G1* incrementa los niveles de expresión GUS en semillas entre 20 y 28 dpa (comparar las fusiones ds10F2 y ds10F2Æ, con ds10F1 en las Figuras 5 y 6A). Esta diferencia es estadísticamente significativa (por ejemplo a 28 dpa: F= 5.397, P=0.0213), y está causada por las

secuencias de *Ha ds10 G1* presentes en la fusión ds10F2Æ (ver Figura 5); ya que no se encontraron diferencias significativas entre la actividad GUS de ds10F2 y ds10F2Æ (por ejemplo, también a los 28 dpa, F=0.274, P=0.6015; ver Figura 6A). En el caso de ds10F2Æ, el efecto estimulador de las secuencias 3'-flanqueantes también se produce, y es altamente significativo, en etapas mas tempranas de la maduración embrionaria (Figura 6A, 16 dpa; F=16.607, P=0.001). En cambio, en estas etapas (entre 12 y 16 dpa) las actividades GUS de ds10F1 y ds10F2 no difieren significativamente entre sí (por ejemplo, a 16 dpa: F=2.762, P= 0.0983; ver Figura 6A). En conjunto estos resultados muestran que ds10F2Æ es la fusión construida y ensayada que funciona mejor en semillas de tabaco desde los 16dpa; y que esto se debe al efecto de las sequencias 3'-flanqueantes de *Ha ds10 G1* incluidas en ella. Desconocemos si este efecto se produce por mecanismos de activación transcripcional, estabilización de ARNm, o por combinación de ambos tipos de mecanismos. En cualquier caso el efecto es claro, y de posible útilidad para diseñar nuevos genes quiméricos de expresión mas eficiente en semillas, desde etapas relativamente tempranas de la maduración embrionaria (veáse también el apartado de "Otros Ejemplos").

Por otra parte, la comparación entre las actividades GUS de las plantas con las fusiones ds10F1 y ds10F3 nos permitió investigar los posibles efectos de la presencia del intrón (y/o de las secuencias codificantes de *Ha ds10 G1* en las que difieren estas fusiones, Figura 5) sobre la expresión de ambas. En semillas de tabaco transgénico estas comparaciones demuestran que la presencia del intrón (mas el primer exón completo y parte del segundo exón) no tiene efectos positivos sobre la expresión GUS, que por lo tanto debe de estar básicamente conferidas por el promotor y secuencias de *Ha ds10 G1* presentes en ds10F1 (Figura 6A). Así por ejemplo, las actividades de ds10F1 y ds10F3 no difieren estadísticamente entre 12 y 28 dpa, salvo a los 20 dpa (F= 4.73, P=0.031), y entonces la presencia de las secuencias adicionales en ds10F3 r dujo significativamente la actividad GUS observada. Por lo tanto, aunque es altamente probable qu el intrón se procese corr ctamente en semillas d sistemas heterólogos como el tabaco (carecemos de una prueba formal de ello),

su posible papel regulador en el desarrollo embrionario no está claro. Sin embargo otras observaciones no excluyen que las secuencias adicionales de *Ha ds10 G1* en ds10F3 (incluyendo el intrón) puedan tener papeles reguladores en otros tejidos (ver, mas adelante, el efecto de éstas secuencias sobre la expresión residual de las fusiones ds10::GUS en el polen y en germínulas).

La especifidad embrionaria (en semillas) de la expresión GUS conferida por las secuencias *Ha ds10 G1* en plantas transgénicas de tabaco se investigó verificándola en otros tejidos; tanto en ausencia de estrés, como tras tratamientos de desecación o con ABA. En el caso de las plantas T0, el único tejido en el que, tanto mediante ensayos fluorimétricos como histoquímicos, se detectó actividad GUS fue en el polen maduro. En otros tejidos las actividades detectadas apenas superaron las del fondo (plantas de tabaco no transformadas). Por ejemplo, en hojas de plantas T0 de unos dos meses de edad: 0-50 pmol MU/ mg x min. Las actividades detectadas en polen son marginales (casi tres órdenes de magnitud inferiores) comparadas con las de semillas de las mismas plantas transgénicas. Además dicha expresión pudiera ser artefactual y depender del uso, como indicador, del gen GUS en las fusiones [según Uknes S, Dincher S, Friedrich L, Negrotto D, Williams S, Thompson-Taylor H, Potter S, Ward E, y Ryals J, en *the Plant Cell* 5: 159-169, 1993]. Sin embargo, de forma sorprendente, observamos que la actividad medida en el polen de 9 plantas ds10F3 fue (136 ±64 pmol MU/ mg x min) significativamente inferior a la de 5 plantas ds10F1 (6427 ±1294 pmol MU/ mg x min; F= 72.573, P= 0.0001). Esto último pudiera indicar que, a diferencia de lo que ocurre en semillas durante la mayor parte de la maduración del embrión (Figura 6A), la presencia de las secuencias adicionales de *Ha ds10 G1* en ds10F3 (incluyendo el intrón) pudiera reducir la expresión, de genes quiméricos que las contengan, en otros tejidos o momentos del desarrollo.

Adicionalmente, se verificó si la expresión de las fusiones ds10::GUS puede inducirse por hormonas (ABA) o tratamientos de estrés (déficit de agua) en plantas transgénicas (T1) de tabaco en distintos momentos de su ciclo vegetativo. Para ello seleccionamos, tras germinación en medio MS con 300 µg/ml de kanamicina, descendientes de 8 plantas originales distintas

conteniendo las fusiones ds10F1, ds10F2Æ y ds10F3; y otras 6 con ds10F2. Las germínulas resistentes se transplantaron a medio MS. Se realizaron distintos experimentos con germínulas, tanto a 8 como a 15 días tras la imbibición. Para los tratamientos con ABA, las germínulas se transplantaron a placas de MS suplementadas con 100 µM ABA y se cultivaron en dicho medio durante 4 días a 25 °C y con iluminación. Las germínulas control también se transplantaron a medio MS sin ABA. El estrés hídrico se provocó colocando a las germínulas durante unas 5-6 horas dentro de una cabina de flujo entre dos papeles de filtro. Tras los distintos tratamientos, las germínulas se procesaron bien individualmente (para los ensayos histoquímicos con X-gluc, mediante incubaciones de 14 h a 25 °C); o conjuntamente (*pool analysis*), para los ensayos fluorimétricos de la actividad GUS, realizados como se ha descrito anteriormente. Los tratamientos de plantas transgénicas adultas, se hicieron usando plantas individuales propagadas como clones vegetativos obtenidos de cada planta original. Para ello, las germínulas seleccionadas de cada planta transgénica se transplantaron a vermiculita embebida con medio Hoagland 0.5X. De cada germínula se obtuvieron tres explantes completos, que tras recuperarse se pusieron en cultivo hidropónico en medio Hoagland líquido (0.5X). Los experimentos se realizaron cuando las plantas se habían recuperado por completo del proceso de propagación, y tenían raíz, tallo y unas 10-12 hojas. Por lo tanto, para los distintos tratamientos se usaron plantas idénticas genéticamente procedentes de cada germínula transgénica seleccionada. Los tratamientos con ABA se hicieron añadiendo la hormona al medio (100 µM), analizándose la actividad GUS en las plantas a las 24h. El estrés hídrico se indujo retirando la raíz del contenedor con el medio, analizándose igualmente las plantas a las 24h tras iniciar el tratamiento. El efecto de los distintos tratamientos se analizó en tres experimentos independientes realizados con los siguientes números de plantas T1 para cada fusión (entre paréntesis el número de plantas T0 de las que proceden en cada caso): ds10F1, 11 (6); ds10F2, 10 (5); ds10F2Æ, 5 (3); y ds10F3, 10 (5).

Los experimentos realizados tanto con germínulas como con plantas adultas confirmaron la especificidad mbrionaria de la expresión conferida por

las secuencias de *Ha ds10 G1* a las distintas fusiones; aportando además indicios adicionales sobre el posible papel regulador de las secuencias de *Ha ds10 G1* presentes en ds10F3 (incluyendo el intrón) mencionadas anteriormente. Así, tanto en plantas adultas control, como tratadas, se detectaron actividades GUS mínimas (de 3 a 300 pmol MU/ mg x min) en todos los tejidos analizados (raiz, tallo, hojas y meristemo apical). Estos niveles de actividad están ligeramente por encima de los valores de fondo y pueden detectarse sólo fluorimetricamente (datos no mostrados).

En germínulas de 8 dpi la expresión de todas las fusiones es unos dos órdenes de magitud inferior a los valores máximos alcanzados en semillas. Esta expresión decrece rápidamente entre los 8 y 15 dpi (por ej. ds10F1 pasa de 2864 ±182 a 813±104 pmol MU/ mg x min); y se restringe exclusivamente en los tejidos embrionarios (cotiledones), sin detectarse en otros tejidos vegetativos (radícula, hipocótilo, hojas) diferenciados tras la germinación (Figuras 6D y E, y datos no mostrados para las otras fusiones). Estos resultados confirman, en plantas transgénicas de tabaco, la especificidad embrionaria de la regulación por secuencias de *Ha ds10 G1*. Además de la reducción general de los valores de actividad GUS mencionada anteriormente, se observaron diferencias entre los valores de las distintas fusiones, algunas de ellas estadísticamente significativas. Estas diferencias son similares cualitativamente a las observadas en semilla (Figura 6A). Entre ellas, y por su posible interés aplicado, ilustramos la reducción de la expresión tras la germinación, mediada por las secuencias de *Ha ds10 G1* presentes en ds10F3 (incluyendo el intrón). Este efecto se observa, como una reducción significativa de actividad GUS al comparar los patrones de expresión de plantas ds10F1 y ds10F3 (Figuras 6D y E). El análisis estadístico de los datos cuantitativos de ds10F1 y ds10 F3 confirmó la significancia de esta diferencia, tanto a los 8 dpi (F= 4.36, P= 0.04) como a los 15 dpi (F= 4.39, P= 0.039). Adicionalmente, con las germínulas de ds10F1 se observó a los 8dpi una moderada inducción de GUS por los tratamientos con ABA que es estadísticamente significativa (de 2864 ±182 a 5790 ±733 pmol MU/ mg x min; F= 5.413, P= 0.023). En el caso de ds10F3 no hubo inducción significativa por el mismo tratamiento (de 1502 ±195 a 2338 ±211 pmol MU/ mg x min; F= 2.58, P=

0.11). Los distintos tratamientos no afectaron substancialm nte la especificidad de tejidos, o el orden de magnitud de la expresión observada para las distintas fusiones ds10::GUS (datos no mostrados).

OTROS EJEMPLOS:

Igualmente pueden obtenerse, de forma análoga a la descrita con detalle en el ejemplo anterior, otros genes quiméricos que contengan secuencias 5'-flanqueantes, y(o) 3'-flanqueantes (terminadores), y(o) codificantes, procedentes *del Ha ds10 G1*, combinadas con secuencias procedentes de otros genes. Estos ejemplos no suponen complicaciones técnicas adicionales a los descritos con mas detalle en los apartados anteriores, por lo que son fácilmente realizables por personas con conocimientos suficientes en el sector de la técnica de la invención. Así por ejemplo, en las fusiones ds10::GUS, las secuencias *Ha ds10 G1* pudieran haber incluido otras secuencias 5'-flanqueantes (Figura 1) mas largas del mismo gen para aumentar su nivel de expresión en semillas, como describimos por ejemplo en [Coca MA, Almoguera C, Thomas TL, y Jordano J, en *Plant Molecular Biology* , 31: 863-876, 1996]. Igualmente, las secuencias GUS podrían ser substituidas por otras que codifiquen distintas proteínas o péptidos (naturales o artificiales), cuya producción regulada en semillas de plantas pudiera ser de interés industrial. Ejemplos de estas últimas posibilidades, dados de forma no exclusiva, serían la fusión a secuencias de *Ha ds10 G1* de secuencias codificantes de genes implicados en la biosíntesis de ácidos grasos en semillas [Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke SE y Davies HM, en *Science*, 257:72-74, 1992], de proteínas de reserva con composiciones ricas en determinados aminoácidos [Saalbach I, Pickardt T, Machemehl F, Saalbach G, Schieder O, y Muntz K, en *Molecular and General Genetics* 242: 226-236, 1994], o de péptidos con actividades antigénicas o farmacológicas [Vandekerckhove J, Van Damme J, Van Lijsebettens M, Botterman J, De Block M, Vandewiele M, De Clercq, Leemans J Van Montagu, M y Krebbers E, en *BioTechnology* 7: 929-932, 1989]. Estas fusiones se realizarían y utilizarían de forma análoga a como se describe en las publicaciones citadas a título de ej mplo (dados de forma no excluyente)

en cada caso. Para facilitar estas posibilidades, hemos construido un plásmido (ds10EC1) qu contiene una *cassette* de expresión que incluye el promotor y las secuencias 5'- y 3'- flanqueantes de *Ha ds10 G1* presentes en ds10F2Æ (ver Figura 5). Entre ambas secuencias y mediante mutagénesis dirigida [Chen E y Przybila AE, en *BioThecniques* 17: 657-659, 1994] hemos añadido un sitio de restricción de Eco RI, que permite la inserción de secuencias de genes, o correspondientes a péptidos, como los mencionados anteriormente (disponibles en otros laboratorios, o que pudieran diseñarse o sintetizarse). El plásmido ds10EC1 se construyó a partir de ds10G1S3Æ10.5 (Figura 1). A partir de dicho plásmido, amplificamos por PCR las secuencias de *Ha ds10 G1* entre las posiciones -1574 (Sal I) y +98; usando ADN polimerasa Pfu y los cebadores 5'-ATTAACCCTCACTAAAG-3' (T3) y 5'-GAGTGAACAgAATtcCATCACAACAGGG-3' (ds10Eco RI). En este último los tres cambios de secuencia (señalados en minúscula) introducen el nuevo sitio de Eco RI en la posición del codón de iniciación. Tras la PCR se purifica un fragmento de ADN de 199 pb (*megaprimer*), que junto con el cebador 5'-AATACGACTCACTATAG-3' (T7) se usa para una segunda amplificacición por PCR de ds10G1S3Æ10.5. El ADN amplificado (795 pb) se digirió con Eco RI y Sph I. El fragmento de ADN resultante (125 pb), con las secuencias de *Ha ds10 G1* entre Sph I (-126) y el nuevo sitio de Eco RI, se purificó y ligó; reemplazando en ds10G1S3 las secuencias de *Ha ds10 G1* (Figura 1) entre las posiciones -126 (Sph I) y 1086 (Eco RI). Tras este paso, la secuencia amplificada por PCR se verificó mediante secuenciación (método de Sanger) usando el cebador T3. Finalmente se insertó en el plásmido obtenido en el paso anterior un fragmento del ADN genómico de *Ha ds10 G1* (Figura 1), con secuencias entre +1086 (Eco RI) y Å+3000 (Xba I), obteniéndose la *cassette* ds10EC1 (Figura 4), clonada en el plásmido pBluescript SK+. El extremo 3' del ADN de ds10EC1 difiere del de ds10F2Æ únicamente en 119 nucleótidos adicionales, correspondientes a secuencias del intrón y del segundo exón de *Ha ds10 G1* . Además, las secuencias de *Ha ds10 G1* en ds10EC1 difieren d las correspondientes en ds10F2Æ en la ausencia de los nucleótidos 1-98 del primer exón (Figura 5).

Dado que la pr sencia de secuencias adicionales de *Ha ds10 G1* en ds10F3 (incluyendo el intrón, el prim r exón y parte del segundo exón) redujo la expresión de este gen quimérico específicamente en tejidos no embrionarios (Ejemplo 3, Figuras 6D-E), es concebible que dichas secuencias pudieran utilizarse para conferir especifidad de semillas a otros genes quiméricos con distintos promotores. El diseño de dichos genes quiméricos no ofrece difcultades técnicas adicionales a las descritas en apartados anteriores: ver por ejemplo los procedimientos detallados para el uso de intrones de plantas con el fin de impedir la expresión de genes quiméricos en *Agrobacterium* [Mankin SL, Allen GC y Thompson WF. *Plant Molecular Biology Reporter* 15: 186-196, 1997]

Los genes quiméricos que contengan secuencias reguladoras de *Ha ds10G1* podrían ser transformados a otras plantas distintas de tabaco (el sistema modelo usado en el ejemplo 3). Entre las mismas hay plantas de gran importancia económica como por ejemplo: el girasol, la soja, la colza, la "canola", el maíz, el trigo, la cebada, el arroz, la "casava", la judía, el cacahuete, etc; cuya transformación genética es posible y está documentada suficientemente en la literatura científica: veáse por ejemplo Lindsey K, Ed. (1993). [*Plant Tissue Culture Manual.* Kluwer Academic Publishers]; y la revisión por Christou [*Trends in Plant Science.* 1: 423- 431, 1996]. Los resultados mostrados en el ejemplo 3 demuestran que, en tabaco, los genes construidos con secuencias reguladoras de *Ha ds10 G1* tienen una elevada actividad desde etapas relativamente tempranas de la maduración embrionaria, manteniendo además la especificidad de semillas característica de la expresión de *Ha ds10 G1* en girasol. Estos resultados podrían obtenerse también con otras plantas, como las mencionadas anteriormente.

DESCRIPCIÓN DE LAS FIGURAS:

Figura 1. Parte superior: mapa de restricción de las secuencias genómicas de *Ha ds10 G1* que flanquean a su región codificante. Las líneas contínuas sobre el mapa indican los distintos fragmentos de ADN genómico que han sido subclonados en el vector pBluescript SK+ (los nombres de los plásmidos respectivos se indican sobre cada fragmento). Los plásmidos

preparados mediante deleciones con Exo III se indican sobre el plásmido de partida (ds10G1S3ÆSacI), indicando n cada caso el extr mo de la deleción. En la parte inferior de la figura se incluye un mapa de restricción detallado de la región cuya secuencia nucleotídica ha sido determinada. La extensión de las distintas reacciones, usadas para ensamblar las secuencias de ambas cadenas de ADN, se indican mediante flechas horizontales (sobre el mapa para la cadena codificante, y bajo el mapa para la no-codificante). Los sitios de iniciación de la transcripción se indican por flechas. Se incluyen barras de escala para ambos mapas.

Figura 2. Implicación funcional de las secuencias RY1 (-129) en la trans-activación del promotor *Ha ds10 G1*. Experimentos de expresión transitoria realizados tras el bombardeo de embriones de girasol con micro-proyectiles cubiertos de ADN. Se representan los resultados de 5 experimentos independientes en los que las distintas mezclas de plásmidos (detalladas en el Ejemplo 1) se bombardearon por quintuplicado en cada experimento. Se representan las medias de las actividades ß-glucoronidasa (GUS) normalizadas con la actividad luciferasa (LUC), así como los errores *standard* (indicados por barras). Clave: F2, pSKds10F2; F2ÆRY1, pSKds10F2ÆRY1; ABI3, muestras con el plásmido efector. Se aprecia una disminución significativa de la actividad relativa GUS/LUC, consecuencia de la mutación en la caja RY1. Las actividades basales para pSKds10F2 (sin incluir el plásmido efector) son del orden de 46±8.

Figura 3. Patrones de acumulación de los ARNm del gen *Ha ds10 G1* en girasol. La autoradiografía mostrada corresponde a ensayos de protección frente a la RNAsa A, tras hibridar una ribosonda del gen con distintas muestras de ARN total. Se observa la acumulación de mensajeros producidos a partir de los sitios de iniciación de la transcripción de *Ha ds10 G1* (como fragmentos protegidos indicados por las flechas num radas). Estos fragmentos se d tectan sólo n embriones (Emb) desde 10 a 20 dpa, y en semillas maduras (25 dpa); p ro no en otras muestras analizadas, como germínulas (Germ) o germínulas tratadas con ABA (Germ + ABA). El carril tRNA corresponde a hibridaciones

control con ARN t de levadura. Se indican con números y flechas las bandas correspondientes a los mRNAs producidos a partir de los distintos sitios se iniciación. El sitio de iniciación número 3 (indicado entre paréntesis) no ha sido confirmado experimentalmente mediante *primer extension*. En el margen izquierdo se incluyen marcadores moleculares de tamaño (pBR322/Hpa III).

Figura 4. Localización de ARNm en secciones de embriones de girasol a los 12 (A y B), 21 (C-E), y 28 dpa (F-H). En cada caso se usaron las siguientes ribosondas : ds10 (-), A, C, F, H; ds10 (+),E, y 18S ARNr, B, D, G. Barras de escala = 500 µm (Savo en F, 125 µm). Parénquima en empalizada= pp. Las flechas señalan el *procambium*.

Figura 5. Mapas de restricción de las fusiones ds10::GUS y de la *cassette* de expresión optimizada ds10 EC1, construídas en los Ejemplos 3 y 4. Mediante distintos sombreados se indican las secuencias de *Ha ds10 G1* y de otros genes contenidas en cada caso. Los sitios de iniciación de la transcripción a partir del promotor de *Ha ds10 G1* están indicados por flechas.

Figura 6. Expresión de las fusiones ds10::GUS en semillas de plantas trasgénicas de tabaco. Panel A: Compendio de todos los datos cuantitativos (determinaciones fluorimétricas). Se muestra el promedio de las actividades GUS observadas en semillas de las plantas transgénicas (T0) y su evolución en distintos momentos del desarrollo embrionario. Los datos correspondientes a cada fusión se indican mediante los símbolos en el inserto de la parte superior izquierda. La barras indican los errores estándard. Paneles B-E: selección representativa con resultados de los experimentos de localización histoquímica de la actividad GUS: B.- embriones a los 12 dpa (plantas ds10F2Æ, T0). C.- embriones y endospermo a los 16 dpa (plantas ds10F2Æ, T0). D.- germínulas a los 15 dpi en condiciones control (plantas ds10F1, T1). E.- germínulas a los 15 dpi en condiciones control (plantas ds10F3, T1). En los paneles D y E, las flechas señalan los tejidos v getativos sin actividad GUS (hojas e hipocótilo).

REIVINDICACIONES

1 - La secuencia de nucleótidos del gen de girasol *Ha ds10 G1*, incluyendo su promotor y elementos reguladores específicos de semillas, descritos por la SEQ Nº 1, y por los mapas de restricción en la Figura 1; y caracterizados en los Ejemplos 1-3.

2.- Las secuencias, o parte de ellas, idénticas u homólogas (al menos en un 70%, por ejemplo en un 80% y particularmente al menos en un 95%) a la SEQ Nº1 o a su secuencia complementaria.

3.- Genes que contengan las secuencias mencionadas en la reivindicación 1-2 y que se expresen específicamente en semillas, de forma homogénea y abundante, desde etapas tempranas de la maduración. Estos genes pueden construirse y usarse mediante técnicas de ADN recombinante, según detalles en las siguientes reivindicaciones (3-6):

4.- El uso para conferir expresión específica en semillas, mediante técnicas de ADN recombinante, del promotor y secuencias 5'-flanqueantes y codificantes de *Ha ds10 G1* (o de parte de dichas secuencias), contenidas en las construcciones: ds10F1, ds10F2 ds10F2Æ, ds10F3 y ds10EC1 (descritas en la Figura 5).

5.- El uso de las secuencias codificantes y 3´-flanqueantes de *Ha ds10 G1* (o de parte de dichas secuencias), contenidas en las construcciones ds10F2 y ds10F2Æ, para incrementar la expresión de genes quiméricos específicamente en semillas de plantas transgénicas.

6.- El uso de las secuencias codificantes y del intrón de *Ha ds10 G1* (o de parte de dichas secuencias), contenidas en la construcción ds10F3, para incrementar la expresión de otros genes quiméricos en semillas, y(o) para reducirla en otros tejidos; aumentando así la eficiencia y especifidad en semillas de estos genes quiméricos.

7.- Añádase a lo anterior: semilla, parte de semilla y extracto de semilla.

8.- *Cassette* de expresión que contenga una secuencia descrita en las reivindicaciones 1 a 6.

9.- Vector(es) que contenga(n) una secuencia descrita en las reivindicaciones 1 a 7.

10.- Células hospedadoras que contengan una secuencia descrita en las reivindicaciones 1 a 7.

11.- El proceso de obtención de plantas transgénicas caracterizadas por la transformación de una planta (por ejemplo el girasol, la soja, la colza, la *canola,* el maiz, el trigo, la cebada, el arroz, la judía, la *casava,* el cacahuete, el tabaco, etc.), con una *cassette* de expresión descrita en la reivindicación 8.

12.- Procedimientos de producción, por ejemplo de aceite, proteínas, o de substancias bio-activas, usando plantas transgénicas como las descritas en la reivindicación 11.

13.- Productos, por ejemplo aceite, proteínas, o substancias bio-activas, obtenidos según la reivindicación 12.

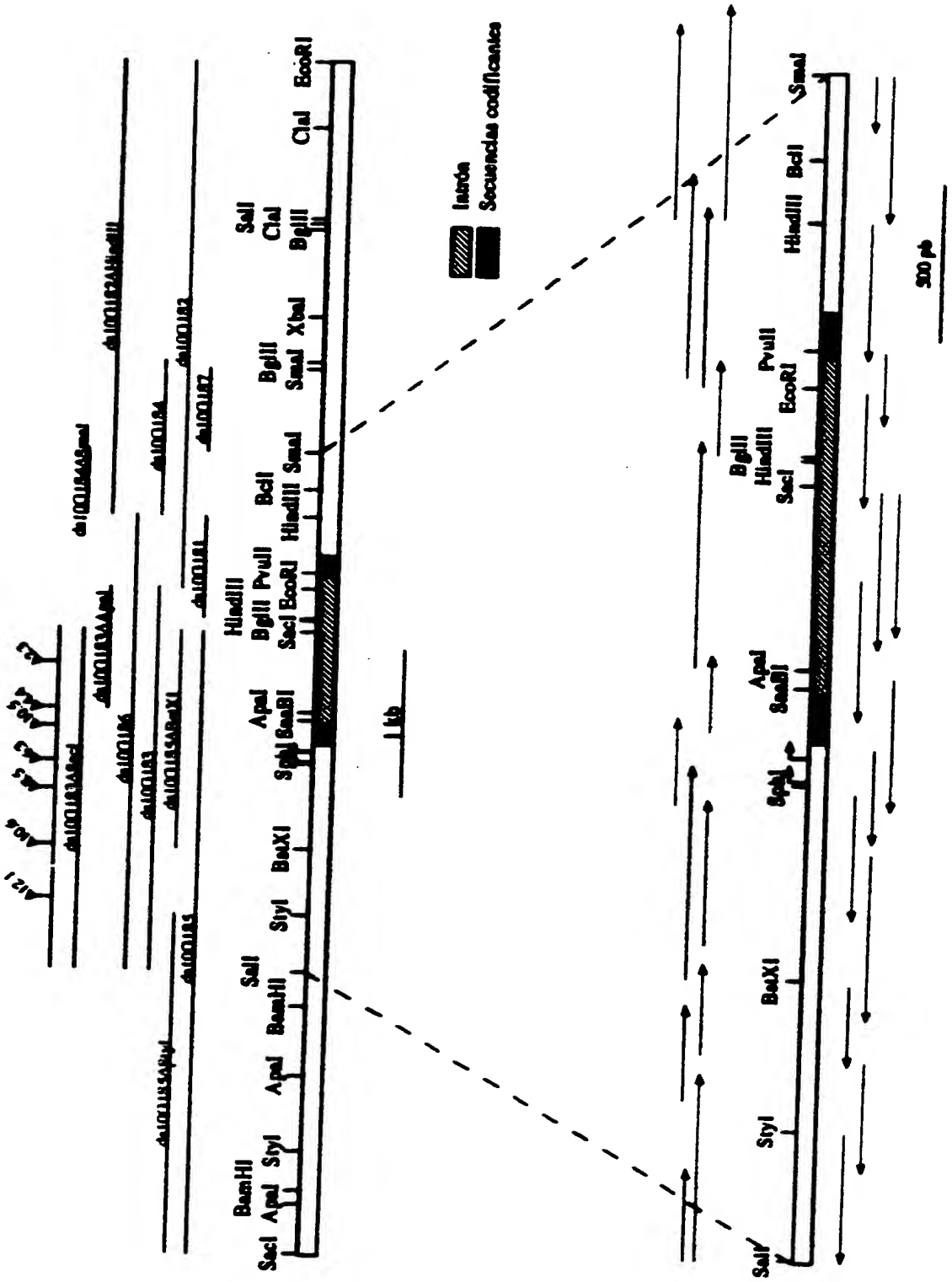

Intrón
Secuencias codificantes
1 kb
500 pb
SacI
ApaI
BamHI
StyI
SalI
BamHI
ApaI
StyI
BclXI
SphI
BamBI
ApaI
HindIII
BglII
PvuII
SacI
BcoRI
HindIII
BclI
SmaI
BglII
SmaI
XbaI
SalI
ClaI
BglII
ClaI
BcoRI
SmaI
BclI
HindIII
PvuII
BcoRI
BglII
HindIII
SacI
ApaI
BamBI
SphI
BclXI
StyI
SalI

Figura 1

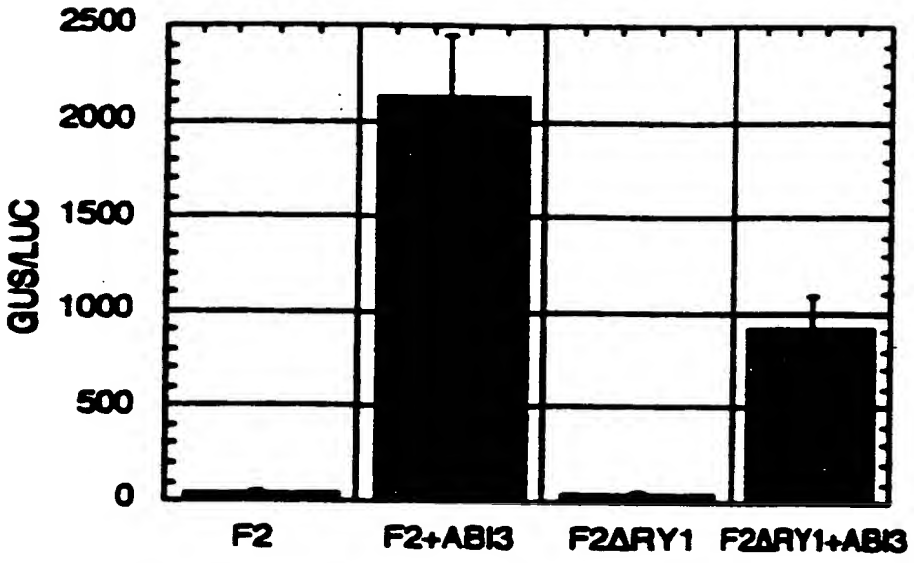


Figura 2

Figura 3

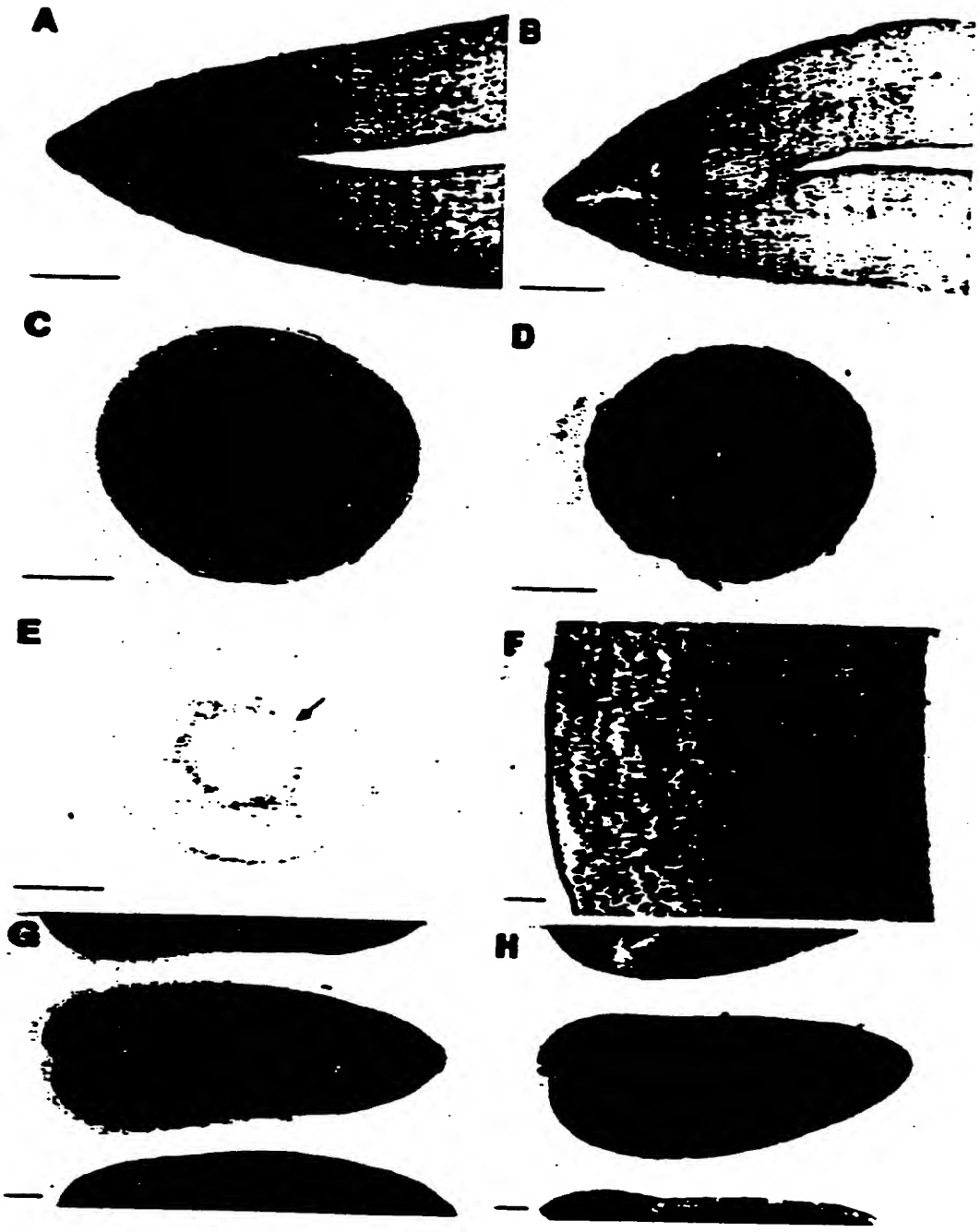


Figura 4

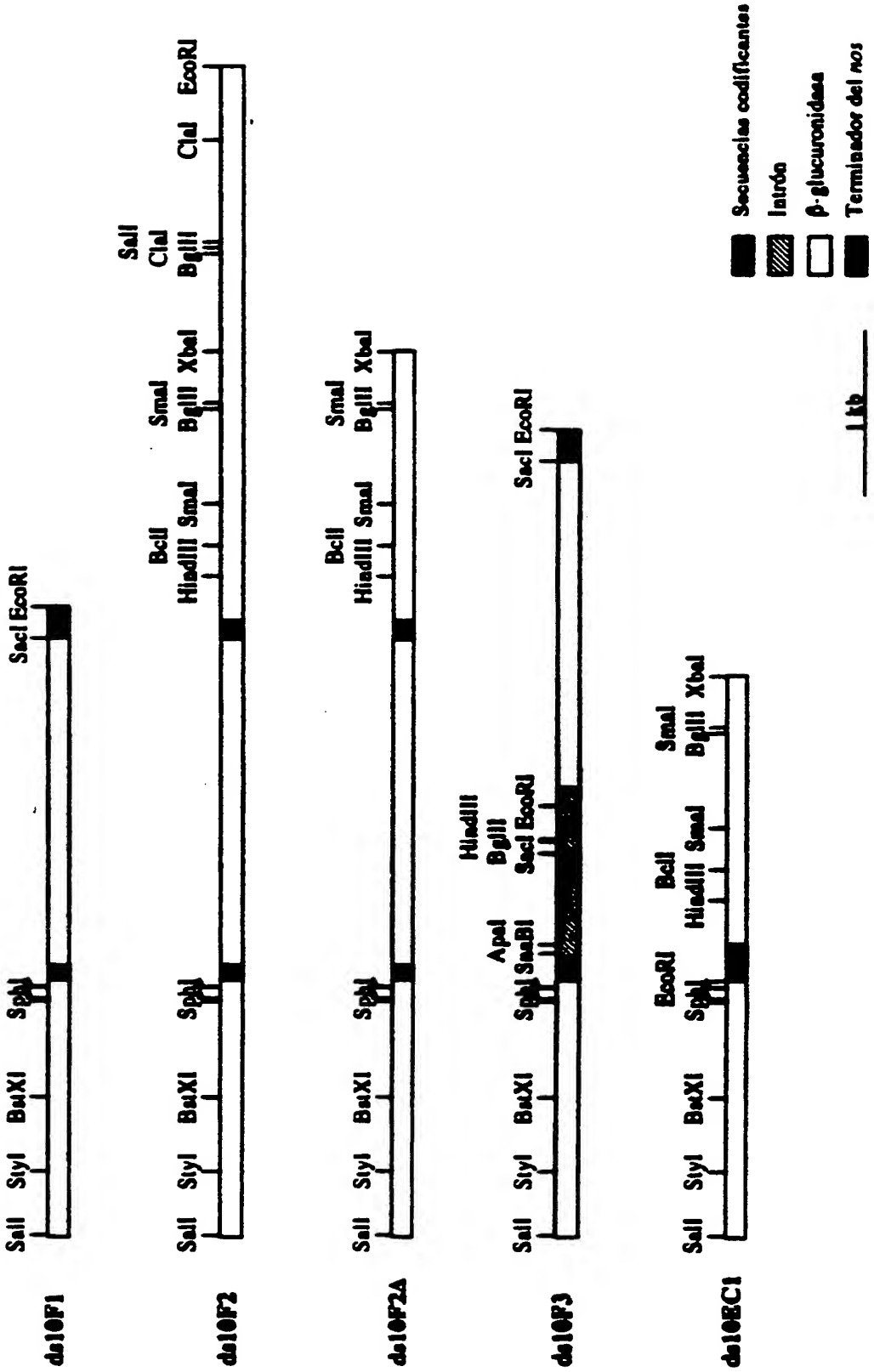
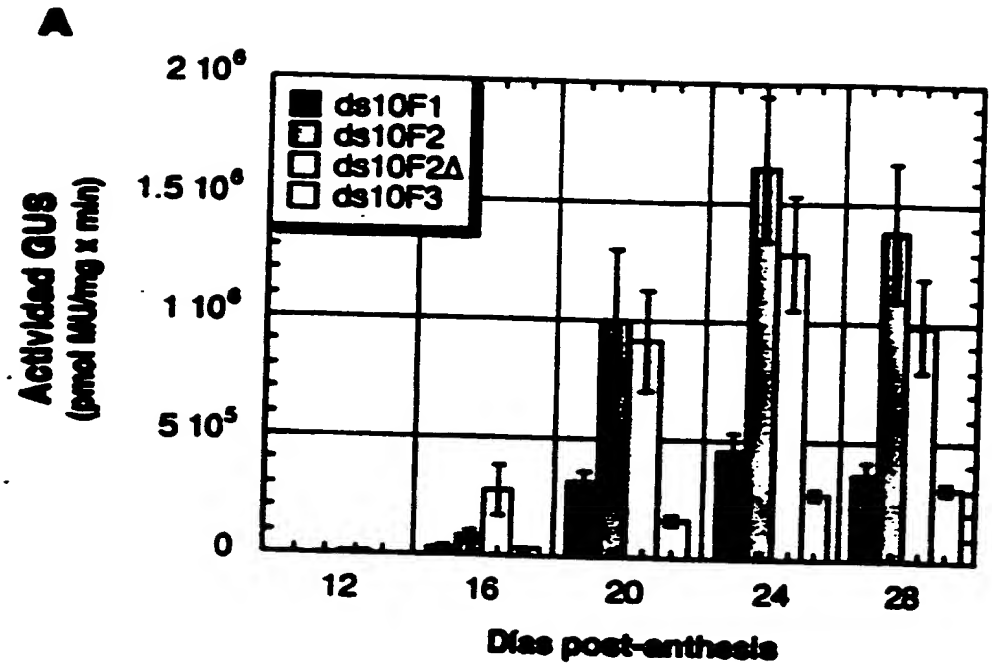

Secuencias codificantes
Intrón
β-glucuronidasa
Terminador del nos
1 kb
pde10F1
pde10F2
pde10F2A
pde10F3
pde10EC1
SacI EcoRI
SalI
StyI
BsuXI
SphI
BclI
HindIII
SmaI
BglII
XbaI
ClaI
ApaI
SnaBI
BamHI

Figura 5


A
2 10⁶
1.5 10⁶
1 10⁶
5 10⁵
0
Actividad GUS
(pmol MU/mg x min)
ds10F1
ds10F2
ds10F2Δ
ds10F3
12
16
20
24
28
Días post-anthesis

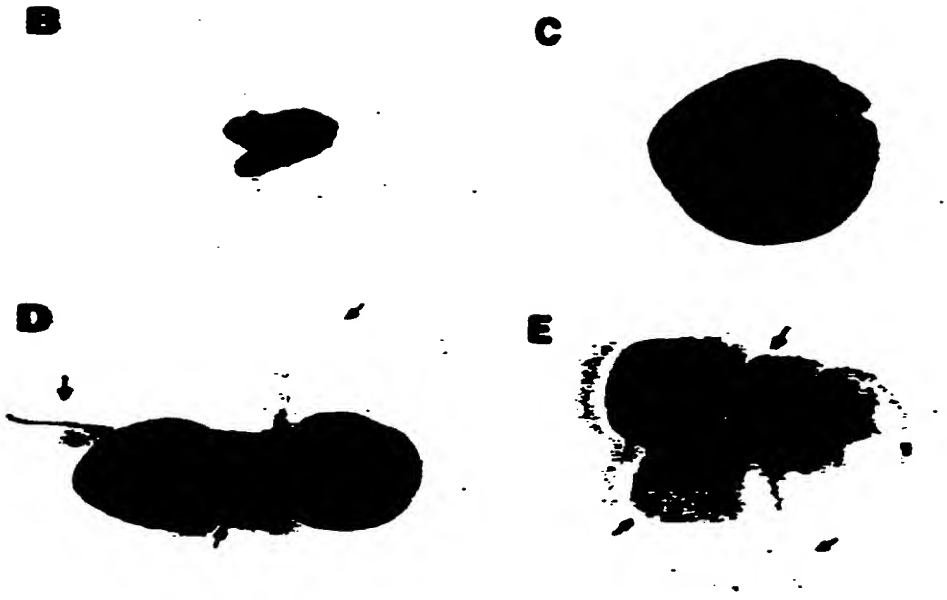

B
C
D
E

Figura 6

LISTA DE SECUENCIAS
SEQ ID No. 1

<110> Consejo Superior de Investigaciones Cientificas

<120> Promotor y secuencias reguladoras de Ha ds10G1: un gen LEA expresado exclusivamente en semillas

<130> Ha ds10G1

<140>
<141>

<160> 1

<170> PatentIn Ver. 2.1

<210> 1
<211> 3617
<212> ADN
<213> Helianthus annuus

<220>
<221> exon
<222> (1577)..(1721)

<220>
<221> exon
<222> (2746)..(2879)

<220>
<221> intron
<222> (1722)..(2745)

<220>
<221> misc_feature
<222> (1448)..(1455)
<223> RY box

<220>
<221> misc_feature
<222> (1512)..(1519)
<223> RY box

<220>
<221> polyA_signal
<222> (3139)..(3144)

<220>
<221> polyA_signal
<222> (3220)..(3225)

<400> 1
gtcgacttct tcatcttcgt ctaagtgttg agtatcgagt acaaattatt catcttcgtt 60
gtcatcgtat atgagaagca tgtgtttaac attatctctt ggatattgag acggtggact 120
ccgataagca aacgggtcaa aagcgttttg ggcttccgta ggatactcgt acacacccgg 180
tgcaaccatc gtgggacttt gtggtaatga aaaaaactgt ggtggatagt tgggttgggg 240
ttgggtgttg taaaaacccg tttgtggaaa aaatgtggtg ggaagttggg ttggtacaca 300
atagatccac ctcgtctact ttgcgagccg cgccccctcg cattcgaacc ggaatccgat 360
actttttttc ttttcatgac ccttgttttg tctatccatg gtatgcatct gaattggtaa 420

attgggtgaa aaatggaatg tttaacatgg taaaatggaa tgttataatt ataaaggtat 480
ttaatgtttt tttttaaacc ataaacggtc atatagccgt ttaaagcaca acggtcaagt 540
ccccaacggt caaatcaagc caaacaatca agtcccgcat gtggcatata tcccctgctt 600
tgaaccaggg ccggcccaag cctaagaatt ttgaggcttg ggctttaggc ctctaaatca 660
atagggactc taacttaaaa aaattatata tgatatttag gttaagttga atttatcttt 720
atatgtacaa aaaaatatat ataattccaa aataaaaaat aaaaaaataa attaacttcg 780
ccaacaataa atcttttgtt atgttttgcc ttttgttttt ttaaaaaaag gctcaaattt 840
ttaatttgct ttaggccacc aaaatggtta aaccgacctt gctttgaacc atacccacac 900
gacaaattag gggatacgga gtgggggtcg gcaaaccaat gccaataaac tttgccgatc 960
aaagtttcac catttcggtg attgtttgcc gatgcggcga ggaggaagaa gagagaggag 1020
ggtagtgtga ggtggggtcc attccaatct caactaatca catttttttc ctttttttta 1080
tttaaaaagt ttaccacttc actaaatgtc taaaccattg ccaacacttt tcaccaaagt 1140
ttaaacattt ttctctgatt gacgtggcac actctcattg gttgattttt tagtttgcca 1200
ctctcaattg tttaaccact ccttacaccc tcttatgtga gcggtggtgt tccccaagcg 1260
acaaaagggc ttaccgcaac cccttaccgc ttccaacctt tacaccctta tcttctgcct 1320
atactgcatg tcactctatt gcagactatc tgagatagct acaacctaac gacttaaggt 1380
gaggacacgt gtatctccaa aacccacttc gtcaccctta ccaccacgtc atcataccac 1440
gtgccaacat gcatgctagt atcttctcta ctatacacat acttatgtac tatatatatt 1500
caccaaaatt acatgcatgt tagcaccctc aattgtacaa cacaacttaa caacttaagt 1560
taaaagagtg aacaaaatgg catcacaaca gggacaacaa acacgcaaaa ttccggaaca 1620
ggagaagaag gatctcgacc aacgagcagc taaaggcgag accgttgttc cgggtggtac 1680
tcgtggcaaa tctcttgagg ctcaagaacg tcttgctgaa ggtatatgct tatttagtta 1740
aatttacata cgtattgggg atggccgttt gatatttagc aaacggcaac ccgaatattt 1800
aggggcccat gatttgacat acatatttga ataagattaa ggggcctaga ggtgtacaaa 1860
tccctttaaa aaagggtccg gttttaaccg gacggaaatg gtttctaaaa tagcggttcg 1920
gtttttaact ggttttacat ggtgggcagt cgaactggta ctagtagatt cggtaagcga 1980
tttgggacaa aaaacgtttt tttaactggt tttgaagatc agtaatacta accggtatgc 2040
aaactgccag ttcgggttga aactggtttc tttttctttt tcaaaaaaga aacggtatgg 2100
gacaaaaaac gttttttttaa actggttttg aagatcaata atactaaccg gtatgcaaac 2160
cgccagttcg gttcgggtta aaactggttt ttttttttt tttaaagaag cggtaaaaaa 2220
accggtttcg gtcataaacc gatttttta acacctacaa gggggccatg atttaaaaaa 2280
aaagcggtaa aaaaaaaccg gtttcggtca taaaccgatt tttttaacac ctacaagggg 2340
gccatgattt aagatgaagt atcggagctc gtgatttaac ataaagtacc tcaaacggcc 2400
gcagtttgat tcaagcgcaa acccgtcgtt ctacgactag tggcgaagct tgagatctcc 2460
gattgggggt cgaaaacgtt tatacccaaa aatttctata gaacgggggt cgaaaacgtg 2520
tatactcaaa aaattctata cgaaagctac atacctgagc gaaaagttcg ggtctggcac 2580
acacccccctc ccccctcttc taacctacgc ccctagctat gacggtgtgg gtagagtcca 2640
ccctttcagt agctttttg cgaattcaca ttagtttatt ttatagttgt agtgatgcat 2700
aataatatat gcatgtactt aattttgtgt ttggtggtgg tgaagggcgg agcaagggag 2760
gacaaacgag gaaagaccag ctgggaactg aagggtacaa ggagatgggg aaaaagggcg 2820
gtcagaccac cggtgacaag tcggctggtg agcgagagga ggaggaggag gaggactaga 2880
tagtcaatag tggtgtgatg gtgtttgcat gtacgatgat gttaatttcc atgttttata 2940
tatgtgtatg tacctgtagt atggtttagc tcgtgtttca tgttttgttg gtcgttttgg 3000
tatcttcttt agtgcatgta cgactagtag tcctatgatg atgtgatgtg atttgcatat 3060
gttgatataa tggtaagtac aagttttatc tatctatctt ttctttcggc tagagtttac 3120
aaaatgtgta cggatgacaa taataagcaa cttatatttt tccatacaag cttattatgt 3180
gactattgcc caataaggca ataagtaagt gagccaataa ataatagccg gttatgctgt 3240
aacaactatt gtgacaacaa tattctttta ccgaataaat aaaaataagt aagtattagg 3300
agccaataaa taaataaccg gtttaatgag aaaagcacca aaatagcctc tttgatcaat 3360
agaattatga aaatagtctc gactcgacct ttttgaataa cccttcatct cctgtgttgc 3420
gcgatcttcg cccggacgac acgtgtctag tgaacatgcc tattctgatt agtggtgcac 3480
aaatcgggtt ttttttttaaa aaacgggttt cgggtaggat cgggtttttt taaaaccggg 3540
tcgggtccgg gtcgggtttt ctaccaaaaa tgtataccct gtataccggg ttcgggtcgg 3600
gttttataaa acccggg 3617

INTERNATIONAL SEARCH REPORT

International application No.
PCT/ ES 99/00017

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 : C12N 15/82, C12N 15/29, A01H 5/00
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 : C12N, A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
CAS, WPI, EPODOC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ALMOGUERA et al. "Developmental and environmental concurrent expression of sunflower dry-seed-stored low-molecular-weight heat-shock protein and Lea mRNAs". 1992 Plant Mol. Biol. Vol. 19(5). Pages 781-92	1-3
A	WO 9713843 A (CORNELL RESEARCH FOUNDATION INC.) 17 April 1997 (17.04.97), page 3, line 13 – page 7, line 25	1-13
A	HULL, G. et al "Analysis of the promoter of an abscisic acid responsive late embryogenesis abundant gene of Arabidopsis thaliana" 1996. Plant Sci. Vol. 114(2). pages 181-92	1-13

☐ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 May 1999 (19.05.99)	2 June 1999 (02.06.99)
Name and mailing address of the ISA/ S.P.T.O	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/ ES 99/00017

Patent document cited in search report	Publication date	Patent familiy member(s)	Publication date
WO 9713843 A	17.04.1997	CA 2234168 A	17.04.1997
		EP 0874897 A	04.11.1998
		AU 7397796 A	30.04.1997

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional nº
PCT/ ES 99/00017

A. CLASIFICACIÓN DEL OBJETO DE LA SOLICITUD

CIP[6] C12N 15/82, C12N 15/29, A01H 5/00
De acuerdo con la Clasificación Internacional de Patentes (CIP) o según la clasificación nacional y la CIP.

B. SECTORES COMPRENDIDOS POR LA BÚSQUEDA

Documentación mínima consultada (sistema de clasificación, seguido de los símbolos de clasificación)

CIP[6] C12N, A01H

Otra documentación consultada, además de la documentación mínima, en la medida en que tales documentos formen parte de los sectores comprendidos por la búsqueda

Bases de datos electrónicas consultadas durante la búsqueda internacional (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

CAS, WPI, EPODOC

C. DOCUMENTOS CONSIDERADOS RELEVANTES

Categoría*	Documentos citados, con indicación, si procede, de las partes relevantes	Relevante para las reivindicaciones nº
X	ALMOGUERA et al. "Developmental and environmental concurrent expression of sunflower dry-seed-stored low-molecular-weight heat-shock protein and Lea mRNAs". 1992 Plant Mol. Biol. Vol. 19(5). Págs. 781-92	1-3
A	WO 9713843 A (CORNELL RESEARCH FOUNDATION INC.) 17.04.1997, pág. 3, línea 13 - pág. 7, línea 25	1-13
A	HULL, G. et al "Analysis of the promoter of an abscisic acid responsive late embryogenesis abundant gene of Arabidopsis thaliana" 1996. Plant Sci. Vol. 114(2). Págs. 181-92	1-13

☐ En la continuación del recuadro C se relacionan otros documentos ☒ Los documentos de familia de patentes se indican en el anexo

* Categorías especiales de documentos citados:

"A" documento que define el estado general de la técnica no considerado como particularmente relevante.

"E" solicitud de patente o patente anterior pero publicada en la fecha de presentación internacional o en fecha posterior.

"L" documento que puede plantear dudas sobre una reivindicación de prioridad o que se cita para determinar la fecha de publicación de otra cita o por una razón especial (como la indicada).

"O" documento que se refiere a una divulgación oral, a una utilización, a una exposición o a cualquier otro medio.

"P" documento publicado antes de la fecha de presentación internacional pero con posterioridad a la fecha de prioridad reivindicada.

"T" documento ulterior publicado con posterioridad a la fecha de presentación internacional o de prioridad que no pertenece al estado de la técnica pertinente pero que se cita por permitir la comprensión del principio o teoría que constituye la base de la invención.

"X" documento particularmente relevante; la invención reivindicada no puede considerarse nueva o que implique una actividad inventiva por referencia al documento aisladamente considerado.

"Y" documento particularmente relevante; la invención reivindicada no puede considerarse que implique una actividad inventiva cuando el documento se asocia a otro u otros documentos de la misma naturaleza, cuya combinación resulta evidente para un experto en la materia.

"&" documento que forma parte de la misma familia de patentes.

Fecha en que se ha concluido efectivamente la búsqueda internacional. 19 Mayo 1999 (19.05.1999)	Fecha de expedición del informe de búsqueda internacional 2 JUN 1999 (02.06.99)
Nombre y dirección postal de la Administración encargada de la búsqueda internacional O.E.P.M. C/Panamá 1, 28071 Madrid, España. nº de fax +34 91 3495304	Funcionario autorizado A. Collados Martín- Posadillo nº de teléfono + 34 91 3495552

Formulario PCT/ISA/210 (segunda hoja) (julio 1998)

INFORME DE BÚSQUEDA INTERNACIONAL

Información relativa a miembros de familias de patentes

Solicitud internacional nº

PCT/ ES 99/00017

Documento de patente citado en el informe de búsqueda	Fecha de publicación	Miembro(s) de la familia de patentes	Fecha de publicación
WO 9713843 A	17.04.1997	CA 2234168 A	17.04.1997
		EP 0874897 A	04.11.1998
		AU 7397796 A	30.04.1997

PCT ORGANIZACION MUNDIAL DE LA PROPIEDAD INTELECTUAL
Oficina Internacional

SOLICITUD INTERNACIONAL PUBLICADA EN VIRTUD DEL TRATADO DE COOPERACION EN MATERIA DE PATENTES (PCT)

(51) Clasificación Internacional de Patentes 6 :		(11) Número de publicación internacional:
C12N 15/82, 15/29, A01H 5/00	A1	WO 99/37795
		(43) Fecha de publicación internacional: 29 de Julio de 1999 (29.07.99)

(21) Solicitud internacional: PCT/ES99/00017

(22) Fecha de la presentación internacional: 23 de Enero de 1999 (23.01.99)

(30) Datos relativos a la prioridad:
P 9800122 23 de Enero de 1998 (23.01.98) ES

(71) Solicitante *(para todos los Estados designados salvo US)*: CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS [ES/ES]; Calle Serrano, 117, E–28006 Madrid (ES).

(72) Inventores; e
(75) Inventores/solicitantes *(sólo US)*: PRIETO–DAPENA, Maria Pilar [ES/ES]; Instituto Recursos Naturales y Agrobiología Sevilla, Consejo Superior de Investigaciones Cientfficas, Apartado 1052 Estafeta–Puerto, E–41080 Sevilla (ES). ALMOGUERA ANTOLINEZ, María Concepción [ES/ES]; Instituto Recursos Naturales y Agrobiología Sevilla, Consejo Superior de Investigaciones Cientfficas, Apartado 1052 Estafeta–Puerto, E–41080 Sevilla (ES). JORDANO FRAGA, Juan Bautista [ES/ES]; Instituto Recursos Naturales y Acrobiología Sevill, a, Consejo Superior de Investigaciones Científicas, Apartado 1052 Estafeta–Puerto, E–41080 Sevilla (ES).

(74) Mandatario: OJEDA GARCIA, Pedro; Consejo Superior de Investigaciones Científicas, Calle Serrano, 113, E–28006 Madrid (ES).

(81) Estados designados: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, Patente ARIPO (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Patente euroasiática (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Patente europea (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), Patente OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Publicada
Con informe de búsqueda internacional.
Con reivindicaciones modificadas.

Fecha de publicación de las reivindicaciones modificadas:
2 de septiembre de 1999 (02.09.99)

(54) Title: PROMOTER AND REGULATOR SEQUENCES *Ha ds10 G1*: A GENE LEA OF SUNFLOWER EXPRESSED EXCLUSIVELY IN SEEDS FROM THE MATURATION PHASE

(54) Título: PROMOTOR Y SECUENCIAS REGULADORAS DE *Ha ds10 G1*: UN GEN LEA DE GIRASOL EXPRESADO EXCLUSIVAMENTE EN SEMILLAS DESDE LA FASE DE MADURACION

(57) Abstract

The present invention discloses the isolation and characterization in transgenic tobacco plants of the promoter and regulator sequences of a gene LEA–I of sunflower, *Ha ds10 G1*. These sequences present characteristics which are extremely appropriate to be used in the modification of seeds (for example of reserve substances). The advantages of their possible use in transgenic plants are demonstrated through examples such as studies related to the accumulation and location of RNAm Ha ds10 in the homologous system. Said studies show both the high expression levels reached during embryogenesis from the early maturation phases and the absolute specificity of the seed, together with a homogenous location in embryos which is finally restricted essentially to the soft tissue in palisade of the cotyledons, a tissue specialized in the accumulation of reserve substances in the sunflower.

(57) Resumen

Con la presente invención aislamos y caracterizamos en plantas transgénicas de tabaco, el promotor y las secuencias reguladoras de un gen LEA–I de girasol, *Ha ds10 G1*. Estas secuencias presentan unas características muy apropiadas para su uso en la modificación de semillas (por ej. de sustancias de reservas). Las ventajas de su posible uso en plantas transgénicas se muestran mediante ejemplos como estudios de la acumulación y localización del ARNm Ha ds10 en el sistema homólogo. Estos estudios muestran tanto los elevados niveles de expresión alcanzados durante la embriogénesis desde fases tempranas de la maduración, como sus absoluta especificidad de semilla, acompañada de una localización homogénea en embriones que acaba restringiéndose fundamentalmente al parénquima en empalizada de los cotiledones, un tejido especializado en la acumulación de sustancias de reservas en el girasol.

UNICAMENTE PARA INFORMACION

Códigos utilizados para identificar a los Estados parte en el PCT en las páginas de portada de los folletos en los cuales se publican las solicitudes internacionales en el marco del PCT.

AL	Albania	**ES**	España	**LS**	Lesotho	**SI**	Eslovenia
AM	Armenia	**FI**	Finlandia	**LT**	Lituania	**SK**	Eslovaquia
AT	Austria	**FR**	Francia	**LU**	Luxemburgo	**SN**	Senegal
AU	Australia	**GA**	Gabón	**LV**	Letonia	**SZ**	Swazilandia
AZ	Azerbaiyán	**GB**	Reino Unido	**MC**	Mónaco	**TD**	Chad
BA	Bosnia y Herzegovina	**GE**	Georgia	**MD**	República de Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tayikistán
BE	Bélgica	**GN**	Guinea	**MK**	Ex República Yugoslava de Macedonia	**TM**	Turkmenistán
BF	Burkina Faso	**GR**	Grecia			**TR**	Turquía
BG	Bulgaria	**HU**	Hungría	**ML**	Malí	**TT**	Trinidad y Tabago
BJ	Benin	**IE**	Irlanda	**MN**	Mongolia	**UA**	Ucrania
BR	Brasil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarús	**IS**	Islandia	**MW**	Malawi	**US**	Estados Unidos de América
CA	Canadá	**IT**	Italia	**MX**	México	**UZ**	Uzbekistán
CF	República Centroafricana	**JP**	Japón	**NE**	Níger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Países Bajos	**YU**	Yugoslavia
CH	Suiza	**KG**	Kirguistán	**NO**	Noruega	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	República Popular Democrática de Corea	**NZ**	Nueva Zelandia		
CM	Camerún			**PL**	Polonia		
CN	China	**KR**	República de Corea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstán	**RO**	Rumania		
CZ	República Checa	**LC**	Santa Lucía	**RU**	Federación de Rusia		
DE	Alemania	**LI**	Liechtenstein	**SD**	Sudán		
DK	Dinamarca	**LK**	Sri Lanka	**SE**	Suecia		
EE	Estonia	**LR**	Liberia	**SG**	Singapur		

REIVINDICACIONES MODIFICADAS
[recibidas por la oficina Internacional el 22 de julio de 1999 (22.07.99);
reivindicaciones 1-13 reemplazadas por las nuevas reivindicaciones 1-25 (2 páginas)]

1.- Secuencia de nucleótidos constituida por el gen de girasol *Ha ds10 G1,* su promotor, secuencias 5'- y 3'-flanqueantes de Ha ds10 G1 y su uso en la expresión génica específica de semilla.

2.- Secuencia de nucleótidos según reivindicación 1, caracterizada porque está constituida por la secuencia SEQ ID Nº1.

3.- Secuencia de nucleótidos según reivindicación 1, caracterizada porque está constituida por fragmentos de la SEQ ID Nº1.

4.- Secuencia de nucleótidos, caracterizada porque comprende una secuencia de nucleótidos según una cualquiera de las reivindicaciones 1 a la 3.

5.- Secuencia de nucleótidos homóloga a la secuencia de nucleótidos según una cualquiera de las reivindicaciones 1 a la 4.

6.- Secuencia de nucleótidos homóloga según reivindicación 5, caracterizada porque la homología es del 70% al 95%.

7.- Secuencia de nucleótidos caracterizada porque contiene secuencias según una cualquiera de las reivindicaciones 1 a la 6 y un gen quimérico.

8.- Secuencia de nucleótidos según reivindicación 7, caracterizada porque permite la expresión de un gen quimérico.

9.- Secuencia de nucleótidos según reivindicación 8, caracterizada porque permite la expresión del gen quimérico especifico de semillas desde etapas tempranas de maduración.

10.- Secuencia de nucleótidos según reivindicación 9, caracterizada porque está constituida por las construcciones ds10F1, ds10F2, ds10F2Æ, ds10F3 y ds10EC1 o parte de dichas secuencias.

11.- Secuencia de nucleótidos según reivindicación 9, caracterizada porque contiene secuencias codificantes y 3'-flanqueantes del gen *Ha ds10 G1.*

12.- Secuencia de nucleótidos según reivindicación 11, caracterizada porque está contenida n las construccion s ds10F2 y ds10F2Æ.

13.- Secuencia de nucleótidos según reivindicación 9, caracterizada porque contien secuencias codificantes y del intrón del gen *Ha ds10 G1.*

14.- Secuencia de nucleótidos según reivindicación 13, caracterizada porque está contenida en la construcción ds10F3.

15.- Cassette de expresión caracterizado porque contiene una secuencia de nucleótidos según una cualquiera de las reivindicaciones 1 a la 14 y un gen quimérico.

16.- Vector caracterizado porque contiene una cassette de expresión según reivindicación 15.

17.- Células hospedadoras caracterizadas porque contienen una secuencia de nucleótidos según una cualquiera de las reivindicaciones 14 a la 16.

18.- Uso de las secuencias de nucleótidos según una cualquiera de las reivindicaciones 1 a la 16 en la expresión específica de genes quiméricos en semillas, parte de semillas, en extracto de semillas, embriones de semillas y en tejidos de germínulas.

19.- Uso de las secuencias de nucleótidos según una cualquiera de las reivindicaciones 10 a la 12 para incrementar la expresión de genes quiméricos específicamente en semillas de plantas transgénicas.

20.- Uso de las secuencias de nucleótidos según una cualquiera de las reivindicaciones 12 a la 14 para incrementar la expresión de genes quiméricos en semillas y/o para reducirla en otros tejidos.

21.- Plantas transgénicas caracterizadas porque son plantas transformadas por una secuencia de nucleótidos según una cualquiera de las reivindicaciones 1 a la 16.

22.- Planta transgénica según reivindicación 21, caracterizada porque se selecciona de las siguientes de girasol, tabaco, soja, colza, la canola, el maíz, el trigo, la cebada, el arroz, la judia, la casava y el cacahuete.

23.- Uso de las plantas transgénicas según una cualquiera de las reivindicaciones 21 a la 22 para la producción de sustancias fruto de la expresión de genes quiméricos.

24.- Uso de las plantas transgénicas según reivindicación 23 caracterizado porque las sustancias son proteinas, substancias bioactivas y aceites.

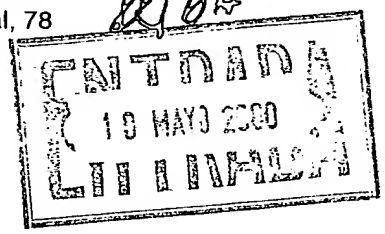
25.- Las sustancias obtenidas según una cualquiera de las reivindicaciones 23 y 24.

ATENT COOPERATION TREAT

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

UNGRIA LOPEZ, Javier
Avda. Ramon y Cajal, 78
28043 Madrid
ESPAGNE



PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL PRELIMINARY EXAMINATION REPORT
(PCT Rule 71.1)

Date of mailing *(day/month/year)*	04.05.00
Applicant's or agent's file reference 199.171/MAD.	**IMPORTANT NOTIFICATION**

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/ES99/00017	23/01/1999	23/01/1998

Applicant
CONSEJO SUPERIOR DE INVESTIGACIONES... et al.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/	Authorized officer
European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Vullo, C Tel.+49 89 2399-8061

Form PCT/IPEA/416 (July 1992)

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
199.171/MAD.		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/ES99/00017	23/01/1999	23/01/1998
International Patent Classification (IPC) or national classification and IPC C12N15/82		
Applicant CONSEJO SUPERIOR DE INVESTIGACIONES... et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 6 sheets, including this cover sheet.

   ☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 3 sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
19/08/1999	04.05.00
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Burkhardt, P Telephone No. +49 89 2399 7456

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/ES99/00017

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-23	as originally filed			

Claims, No.:

1-24	as received on	02/03/2000	with letter of	28/02/2000

Claims, pages:

28-30	as received on	02/03/2000	with letter of	28/02/2000

Drawings, sheets:

1/5-5/5	as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

see separate sheet

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims 1-24
	No:	Claims
Inventive step (IS)	Yes:	Claims
	No:	Claims 1-24
Industrial applicability (IA)	Yes:	Claims 1-24
	No:	Claims

2. Citations and explanations

see separate sheet

Re Item I

Basis of the opinion

The amended claims filed with the letter of 28.02.2000 are formally acceptable under Article 34(2)(b) PCT.

This written opinion is also based on the Sequence Listing (pages 1-2) as filed with the letter of 19.08.1999.

Re Item V

Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

The following document (D) is referred to in this report:

D1 Almoguera and Jordano, 1992. Plant Mol. Biol. 19:781-792.

1. Article 33(2) PCT (Novelty)

1.1 Present claim 1 is directed to the **genomic** sequence of the sunflower *Ha ds10 G1* gene (SEQ ID NO:1). Prior art document D1 discloses the **cDNA** sequence of the sunflower *Ha ds10 G1* gene. Thus, D1 does not anticipate the subject-matter of present claim 1. The same holds true for dependent claim 2 - 16 addressing homologous sequences, expression cassettes, vectors and host cells containing said sequence.

1.2 The use of *Ha ds10 G1* sequences for seed- or seedling-specific expression of chimeric genes in transgenic plants, as laid out in present claims 17 - 19 as well as the resulting plants and the use of these plants (claims 20 - 23) has not yet been disclosed in the prior art presently available to the IPEA.

1.3 For the assessment of novelty of the present "product by process" claim 24 no unified criteria exist in the PCT. The EPO, for example, does not recognize novelty merely by the fact that the product is produced by means of a new

process. Novelty can only be established where use of the method necessarily means that the product has a particular characteristic and that a person skilled in the art following the teaching of the application would inevitably acquire a product which has different characteristics to the product disclosed in the prior art. This does not seem to be the case for present claim 25.

2. Article 33(3) PCT (Inventive step)

2.1 The closest prior art to the subject-matter of present claim 1 appears to be D1. It discloses the cDNA sequence of the sunflower *Ha ds10 G1* gene (page 785, Figure 1) and furthermore states that isolation and characterization of the corresponding genomic sequences will allow further studies on the regulation of the gene (D1, page 790, last paragraph). Claim 1 differs from that in the presentation of the genomic sequence of the sunflower *Ha ds10 G1* gene.

2.2 In the light of the prior art and having regard to the present description and claims, the technical problem may thus be the provision of the genomic *Ha ds10 G1* sequence.
It is common general knowledge in modern biotechnology, and therefore within the scope of a man skilled in the art, to isolate the genomic sequence of a gene for which the cDNA sequence is known. Therefore, the subject-matter of present claim 1 is not based on an inventive concept. The same holds true for present claims 2- 16.

2.3 The expression pattern of the *ds10* gene may be different from that of other *lea* and *lea-a* genes, although the *ds10* expression pattern itself was already known from the prior art (D1, page 787, left column, second paragraph).
Present claim 1, however, is directed to a product, i.e. the complete genomic *ds10* sequence. Its expression pattern is an inherent feature of the promoter sequence. With regard to the technical problem to be solved (see 2.2) this aspect therefore is neglectable.

2.4 A similar objection as in paragraph 2.2 applies to present claim 17 directed to the use of the above nucleotide sequences for specific expression of chimeric

genes in seeds, seed parts, seed extract, seed embryo, and seedling tissue. It has been known from the prior art (D1, page 787, left column, second paragraph) that the transcript of the *Ha ds10 G1* especially accumulates in embryos, dry seeds and seedlings. Therefore, it appears to be obvious to use the corresponding promoter sequences to drive the expression of chimeric genes in the respective tissues. Present claim 18 does not meet the requirements of Article 33(3) PCT. The same holds true for present claims 19 - 24.

CLAIMS

1. A nucleotide sequence constituted by the *Ha ds10 G1* gene, its promoter, Ha ds10 G1 5'- and 3' flanking sequences, wherein the nucleotide sequence is selected from the group consisting of identical nucleotide sequences identical to of SEQ ID NO:1, first homologous nucleotide sequences being homologous by at least 70% to SEQ ID NO:1, second homologous nucleotide sequences being homologous being at least 70% homologous to complementary sequences to SEQ ID NO:1, and fragments thereof.

2. A nucleotide sequence according to claim 1, wherein the first homologous sequence is homologous by at least 80% to SEQ ID NO:1.

3. A nucleotide sequence according to claim 1, wherein the first homologous sequence is homologous by less than 95% to SEQ ID NO:1.

4. A nucleotide sequence, wherein the second homologous sequence is homologous by at least 80% to SEQ ID NO:1.

5. A nucleotide sequence according to claim 1, wherein the second homologous sequence is homologous by less than 95% to SEQ ID NO:1.

6. A nucleotide sequence according to any of the claims 1 to 6, and further including a chimeric gene.

7. A nucleotide sequence according to claim 6, suitable for expression of a chimeric gene.

8. A nucleotide sequence according to claim 7, wherein the chimeric gene is specifical of seeds from early maturation stages.

9. A nucleotide sequence according to claim 8, constituted by constructions ds10F1, ds10F2, ds102Δ, ds10F3 and ds10EC1 or part thereof.

10. A nucleotide sequence according to claim 10, including *Ha ds10 G1* gene coding and 3'-flanking sequences.

11. A nucleotide sequence according to claim 10, including ds10F2 and ds10F2Δ in constructions.

12. A nucleotide sequence according to claim 8, including *Ha ds10 G1* gene coding and intron sequences.

13. A nucleotide sequence according to claim 12, contained in constructions ds10F3.

14. An expression cassette including a nucleotide sequence according to any of claims 1 to 13 and a chimeric gene.

15. A vector including an expression cassette according to claim 14.

16. Host cells including a nucleotide sequence according to any of claims 14 to 15.

17. Use of nucleotide sequences as defined in any of claims 1 to 15, in the specific expression of chimeric genes in seeds, seed parts, seed extract, seed embryos and seedling tissues.

18. Use of nucleotide sequences as defined in any of claims 9 to 11 for increasing the expression of chimeric genes specifically in transgenic plant seeds.

19. Use of nucleotide sequences as defined in any of claims 11 to 13 for increasing the expression of chimeric genes in seeds and/or reduce it in other tissues.

20. A transgenic plant transformed by a nucleotide sequence according to any of claims 1 to 15.

21. A transgenic plant according to claim 20, selected from sunflower, tobacco, soya, oilseed rape, "canola", maize, wheat, barley, rice, bean, cassava

and peanut.

22. Use of a transgenic plant according to any of claims 20 to 21 for the production of substances resulting from the expression of chimeric genes.

23. Use of a transgenic plant according to claim 22 wherein the substances are proteins, bioactive substances and oils.

24. Substances obtained according to any of claims 23 and 24.

UNGRIA®

COMPLETE TEXT INCLUDING ALL THE AMENDMENTS MADE DURING INTERNATIONAL PHASE

UNGRIA®
UNGRIA PATENTES Y MARCAS, S.A.
ORGINA TECNICA INTERNACIONAL DE PATENTES Y MARCAS
FUNDADA EN 1891

EUROPEAN PATENT OFFICE
International Preliminary
Examining Authority (PCT)
D-80298 München
Alemania

28043 - MADRID
Ramón y Cajal, 78
Teléfono: 91 413 60 62*
Telex: 44296 UPAT E
Telefax: 91 413 64 17
e-mail: ungria@ungria.es

08006 - BARCELONA
Tuset, 22
Teléfono: 93 237 50 20
Telefax: 93 237 63 64

46002 - VALENCIA
Lauria, 6
Teléfono: 96 351 35 51
Telefax: 96 352 88 57
e-mail: levante@ungria.es

03002 - ALICANTE
Avda. Explanada, 3
Teléfono: 96 514 12 13
Telefax: 96 520 55 20

48011 - BILBAO
Pl. S. Corazón, 5
Teléfono: 94 441 80 50
Telefax: 94 442 52 17

41018 - SEVILLA
Edificio <<Sevilla 1>>
Teléfono: 95 464 15 93
Telefax: 95 464 88 50

11402 - JEREZ
Cádiz, 5 - Edif. Albaida
Teléfono: 956 18 56 72
Telefax: 956 18 06 30

29001 - MALAGA
Alameda de Colón, 9
Teléfono: 95 222 00 68
Telefax: 95 222 00 54

15005 - A CORUÑA
Enrique Dequidt, 14
Teléfono: 981 14 37 22
Telefax: 981 14 37 30

36201 - VIGO
Colón, 18
Teléfono: 986 22 11 59
Telefax: 986 43 00 58

50004 - ZARAGOZA
Casa Jiménez, 6
Teléfono: 976 21 91 92
Telefax: 976 22 60 27

47001 - VALLADOLID
Duque Victoria, 21
Teléfono: 983 30 64 22
Telefax: 983 39 07 37

UNGRIA INTERNATIONAL, Inc.
15 Essex Road - PARAMUS
NEW JERSEY 07652 (U.S.A.)
Teléfono: 201 368 12 22
Telefax: 201 368 07 63
e-mail: UNGRIAUSA@aol.com

CORRESPONSALES ESPECIALIZADOS EN TODOS LOS PAISES DEL MUNDO

REG. MERC. MADRID, HOJA M-36005, FOLIO 184
TOMO 2.019. INSCRIPCION 1ª. C.I.F. A-28378578

JU/HA/yo **Madrid, 28th February 2000**

By fax: 00/49/89/ 2399-4465 **No. of pages: 5+3+3= 11**

Confirmation by Registered Airmail

Re.: International Patent Application
No.: PCT/ES99/00017
Priority: Spanish Patent Application P9800122, filed on the 23rd January 1998
Applicants: CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
O/ref.: 199.171/MAD

Dear Sirs:

(1) In response to the first Written Opinion issued on the above referenced application, the following is submitted herewith

* replacement sheets 28-30 with amended claims 1-24;

and

* explanatory sheets 28-30 showing the amendments appearing in amended claims 1-24, by bracketed (=deletions) and underlined (=additions) phrases.

As readily apparent, the amended claims are supported in the original disclosure of this application as follows:

Amended claim 1 is supported in modified claim 1, as well as in original claims 1 and 2, and it comprises the reference to SEQ ID NO:1 and does not anymore contain the phrase "use ... in ... gene expression".

Amended claims 2 - 5 are supported in original claim 2.

Modified claim 6 has been deleted.

Amended claims 6-24 are supported in modified claims 7-25.

Furthermore, the amended claims include clerical amendments which are considered to be self-evident and thus admissible.

It is submitted that the amended claims as submitted herewith, are directly and unambiguously derivable from and do not extend beyond the disclosure of the present application as originally filed, and they are thus admissible.

(2) Modified claims 3 and 4 were objected for extending beyond the disclosure of the international application as originally filed.

Whilst claims 3 and 4 have been deleted, amended claim 1 includes the phrase "and fragments thereof" which is supported in line 1 of original claim 2 according to which the claimed sequences "or part of them", were defined as "identical or homologous to" SEQ ID NO:1.

It is submitted that a "part" of a sequence is the same as a "fragment" of a sequence, so that the phrase "and fragments thereof" in amended claim 1 is fully supported in original 2.

It is further submitted that that the phrase "and fragments thereof", is also supported in the description of the present application as originally filed inasmuch chapter "Other Examples" in the description mentions the larger fragments without excluding other possibilities. In view that, as will also be referred to further hereinbelow, the utility of the Class I Lea gene with novel and previously unknown expression patterns has been evidenced and sequence combinations reproducing and improving said expression have been described in the description of the present application, the skilled person would know the feasibility of achieving similar results with fragments of sequence SEQ ID NO:1.

It is submitted that a "part" of a sequence is, in the context of the present invention, the same as a "fragment" of a sequence, so that the phrase "and fragments thereof" in amended claim 1 is fully supported in original 2 and in the description as originally filed and does not extend beyond the original disclosure of this application.

(3) The examiner has objected modified claims 1-15 for lacking an inventive step. In respect hereof, the following is submitted.

(3.1) Whilst applicants agree with the examiner in that the disclosure given by D1 regarding the Ha ds10 cDNA sequence would facilitate the cloning of the

corresponding genomic sequence, the skilled person would not have been able to infer from D1 or from other literature, that the ds10 expression pattern would differ from those of other LEA and LEA-A genes, nor to define regulatory sequences from Ha ds10 cDNA G1 which would be sufficient to reproduce and improve the novel expression patterns by using chimeric genes in transgenic plants. In fact, the expression patterns were not known and, even if the expression patterns and genomic sequence had been known, such a reproduction and improvement would not have been predictable nor obvious on the grounds of the disclosure of D1.

(3.2) Whilst the crucial information the skilled person would not have been able to infer from D1 was obtained experimentally by determining the Ha ds10 G1 novel expression patterns as well as verifying and modifying the expression patterns in transgenic seeds of chimeric genes containing Ha ds10 G1 sequences whereby the genomic sequence could be obtained by using the published cDNA sequence and usual genomic library construction and hybridization techniques, D1 is silent in respect of why this should be useful for seed genetic engineering in view that the expression patterns of Ha ds10 G1 was unknown.

Moreover, on the grounds of the disclosure of D1, the skilled person did not know which of the genomic sequences (or combination of sequences) would be suitable for the construction of chimeric genes with an appropriate expression pattern.

(3.3) The expression pattern of Ha ds10 G1 was, as stated hereinabove, unknown and could not be predicted from the disclosure in D1. In fact, D1 discloses the cDNA nucleotide sequence and nucleic acid (DNA and mRNA) hybridizations using the cDNA as a probe. The hybridization disclosed in D1 does not allow to distinguish the expression pattern of genes being homologous to the cDNA from that of the gene exactly corresponding to the cDNA (Ha ds10 G1).

Thus, table 1 in D1 discloses that the cDNA hybridized, under the experimental conditions disclosed in D1, with 3 to 5 homologous genes, and, in fact, D1 cautiously uses the term "ds10 homologous mRNAs" to describe the observed hybridization patterns.

In connection herewith, it should be noted that D1 did not permit to conclude that the expression of Ha ds10 G1 would be seed-specific inasmuch inducible accumulation of ds10 homologous mRNAs in seedlings (after ABA, mannitol, and heat stress treatments - cf. figures 4 and 5 in D1) was observed. Neither could mRNA accumulation from early stages of seed maturation (cf. fig. 3, 13 daf) be unambigously ascribed to Ha ds10 G1 as D1 does not contain data which would have taught the skilled person which of the 3 to 5 homologous genes would give rise to the observed hybridization signal.

(3.4) Homogeneous accumulation of Ha ds10 mRNA in seeds from early maturation is not only not described in D1 but could neither be predicted from the cDNA and not even from the genomic sequence.

The skilled person would thus not have been able to find, without the need of performing an inventive step, that the Ha ds10 G1 expression patterns would be a combination of an early maturation accumulation (which was known for other for other genes of the LEA I family) **with a homogeneous distribution of mRNA** in immature seeds, and with a change of distribution to mostly the palisade parenchym at later maturation stages. This combination including said distribution which is unique to Ha ds10 G1, could not be predicted and not even suspected on the grounds of available prior art in view that there were no similar precedents in literature, and could not be found until the surprising and unprecedented expression patterns were experimentally determined in the present application.

Without knowing the data disclosed in the present invention and without inventive activity, it would have been impossible to a skilled person to realize or predict, the expression patterns and the so inferred utility of the promoter and regulatory sequences contained in the Ha ds10 G1 sequences.

Further, even after knowing the Ha ds10 G1 expression patterns, the skilled person would not yet have known with which genes and which possible differing combinations of promoter and regulatory sequences, there could be defined combination(s) of sufficient sequences which would reproduce similar expression patterns in transgenic plants. This crucial information would not have been derivable from the genomic sequence as such because the necessary sequences could be located anywhere in the gene, including the 5'-flanking, coding, intron and 3'-flanking region(s), or in different combination(s) of these sequences. In fact, as evidenced by the present application, some of the tested combinations improve the natural expression pattern of Ha ds10 G1, as for example inclusion of the shorter 3'-flanking region sequences in the F2Δ and ds10EC1 chimeric genes which increase the level of expression during early maturation.

(3.5) The unexpected effect of Ha ds10 G1 and of chimeric genes containing promoter and regulatory sequences derived therefrom, is the advantage resulting from the combination of seed-specificity with gene activity from early maturation, and with a homogeneous distribution of gene expression that later evolves to a pallisade parenchyma tissue specificity.

As stated in the present application and set forth hereinabove, it is this unique combination of properties that makes the nucleotide sequences of the present invention especially useful for seed genetic engineering.

UNGRIA®

Whilst the skilled person would have known that there were other examples of seed-specificity and of early expression in the LEA-I gene family, none of such examples was known to be expressed with the tissue specificity of the Ha ds10 G1 of the present invention.

In fact, on the grounds of the more restricted expression patterns for other, similar genes, the skilled person would have *prima facie* inferred that LEA genes in general, despite their specificity and, in some cases as in connection with LEA-A genes, rather early activation, would not be good candidates for seed modification involving gene expression during maturation.

(3.6) In view of the above, it is submitted that the nucleotide sequences of the present invention and the effects being achievable by using said sequences, are not derivable form the disclosure of D1 in an obvious manner and, moreover, lead to an unexpected advantage, and therefore involve an inventive step.

(4) Favorable consideration of the present submissions is earnestly solicited. Should the examiner wish to discuss any issues on the phone, we shall be available at phone # 00/34/91/ 413-6062.

Respectfully submitted,

Javier UNGRIA

Enc.: as mentioned above

CLAIMS

1. A nucleotide sequence constituted by the *Ha ds10 G1* gene, its promoter, Ha ds10 G1 5'- and 3' flanking sequences, wherein the nucleotide sequence is selected from the group consisting of identical nucleotide sequences identical to of SEQ ID NO:1, first homologous nucleotide sequences being homologous by at least 70% to SEQ ID NO:1, second homologous nucleotide sequences being homologous being at least 70% homologous to complementary sequences to SEQ ID NO:1, and fragments thereof.

2. A nucleotide sequence according to claim 1, wherein the first homologous sequence is homologous by at least 80% to SEQ ID NO:1.

3. A nucleotide sequence according to claim 1, wherein the first homologous sequence is homologous by less than 95% to SEQ ID NO:1.

4. A nucleotide sequence, wherein the second homologous sequence is homologous by at least 80% to SEQ ID NO:1.

5. A nucleotide sequence according to claim 1, wherein the second homologous sequence is homologous by less than 95% to SEQ ID NO:1.

6. A nucleotide sequence according to any of the claims 1 to 6, and further including a chimeric gene.

7. A nucleotide sequence according to claim 6, suitable for expression of a chimeric gene.

8. A nucleotide sequence according to claim 7, wherein the chimeric gene is specifical of seeds from early maturation stages.

9. A nucleotide sequence according to claim 8, constituted by constructions ds10F1, ds10F2, ds102Δ, ds10F3 and ds10EC1 or part thereof.

10. A nucleotide sequence according to claim 10, including *Ha ds10 G1* gene coding and 3'-flanking sequences.

11. A nucleotide sequence according to claim 10, including ds10F2 and ds10F2Δ in constructions.

12. A nucleotide sequence according to claim 8, including *Ha ds10 G1* gene coding and intron sequences.

13. A nucleotide sequence according to claim 12, contained in constructions ds10F3.

14. An expression cassette including a nucleotide sequence according to any of claims 1 to 13 and a chimeric gene.

15. A vector including an expression cassette according to claim 14.

16. Host cells including a nucleotide sequence according to any of claims 14 to 15.

17. Use of nucleotide sequences as defined in any of claims 1 to 15, in the specific expression of chimeric genes in seeds, seed parts, seed extract, seed embryos and seedling tissues.

18. Use of nucleotide sequences as defined in any of claims 9 to 11 for increasing the expression of chimeric genes specifically in transgenic plant seeds.

19. Use of nucleotide sequences as defined in any of claims 11 to 13 for increasing the expression of chimeric genes in seeds and/or reduce it in other tissues.

20. A transgenic plant transformed by a nucleotide sequence according to any of claims 1 to 15.

21. A transgenic plant according to claim 20, selected from sunflower, tobacco, soya, oilseed rape, "canola", maize, wheat, barley, rice, bean, cassava

and peanut.

22. Use of a transgenic plant according to any of claims 20 to 21 for the production of substances resulting from the expression of chimeric genes.

23. Use of a transgenic plant according to claim 22 wherein the substances are proteins, bioactive substances and oils.

24. Substances obtained according to any of claims 23 and 24.

CLAIMS

1.[-] A [N]nucleotide sequence constituted by the *Ha ds10 G1* gene, its promoter, Ha ds10 G1 5'- and 3' flanking sequences [and uses thereof in specific gene expression in seeds], wherein the nucleotide sequence is selected from the group consisting of identical nucleotide sequences identical to of SEQ ID NO:1, first homologous nucleotide sequences being homologous by at least 70% to SEQ ID NO:1, second homologous nucleotide sequences being homologous being at least 70% homologous to complementary sequences to SEQ ID NO:1, and fragments thereof.

2.[-] A [N]nucleotide sequence according to claim 1, [characterized in that it is comprised of the] wherein the first homologous sequence is homologous by at least 80% to SEQ ID [No.] NO:1.

3.[-] A [N]nucleotide sequence according to claim 1, [characterized in that it is comprised of fragments of] wherein the first homologous sequence is homologous by less than 95% to SEQ ID [N°]NO:1.

4.[-] A [N]nucleotide sequence[, characterized in that it comprises a nucleotide sequence] according to [any of] claim[s] 1 [to 3], wherein the second homologous sequence is homologous by at least 80% to SEQ ID NO:1.

5.[-] A [N]nucleotide sequence [homologous to the nucleotide sequence] according to [any of the] claim[s] 1 [to 4]], wherein the second homologous sequence is homologous by less than 95% to SEQ ID NO:1.

[6.- Homologous nucleotide sequence according to claim 5, characterized in that the homology is from 70% to 95%.]

[7.-] 6. A [N]nucleotide sequence [characterized in that it contains sequences] according to any of the claims 1 to 6, and further including a chimeric gene.

[8.-] 7. A [N]nucleotide sequence[,] according to claim [7] 6, [characterized in that it permits the] suitable for expression of a chimeric gene.

[9.-] 8. A [N]nucleotide sequence[,] according to claim [8] 7, [characterized in that it permits the expression of] wherein the chimeric gene is specifical[ly] of seeds from early maturation stages.

[10.-] 9. A [N]nucleotide sequence[,] according to claim [9] 8, [characterized in that it is] constituted by [the] constructions ds10F1, ds10F2, ds102Δ, ds10F3 and ds10EC1 or part [of said sequences] thereof.

[11.-] 10. A [N]nucleotide sequence[,] according to claim [9] 10, [characterized in that it contains] including *Ha ds10 G1* gene coding and 3'-flanking sequences.

[12.-] 11. A [N]nucleotide sequence[,] according to claim [11] 10, [characterized in that it contains in the constructions] including ds10F2 and ds10F2Δ in constructions.

[13.-] 12. A [N]nucleotide sequence[,] according to claim [9] 8, [characterized in that it contains] including *Ha ds10 G1* gene coding and intron sequences.

[14.-] 13. A [N]nucleotide sequence[,] according to claim [13] 12, [characterized in that it is] contained in [the] constructions ds10F3.

[15.-] 14. An [E]expression cassette [characterized in that it contains] including a nucleotide sequence according to any of [the] claims 1 to [14] 13 and a chimeric gene.

[16.-] 15. A [V]vector [characterized in that it contains] including an expression cassette according to claim [15] 14.

[17.-] 16. Host cells [characterized in that it contains] including a nucleotide sequence according to any of [the] claims [13] 14 to [16] 15.

[18.-] 17. Use of nucleotide sequences [according to] as defined in any of [the] claims 1 to [16] 15, in the specific expression of chimeric genes in seeds, seed parts, seed extract, seed embryos and seedling tissues.

[19.-] 18. Use of [the] nucleotide sequences [according to] as defined in any of [the] claims [10] 9 to [12] 11 [in order to increase] for increasing the expression of chimeric genes specifically in transgenic plant seeds.

[20.-] 19. Use of nucleotide sequences [according to] as defined in any of [the] claims [12] 11 to [14] 13 [in order to increase] for increasing the expression of chimeric genes in seeds and/or reduce it in other tissues.

[21.-] 20. A [T]transgenic plant[s] [characterized in that they are plants] transformed by a nucleotide sequence according to any of claims 1 to [16] 15.

[22.-] 21. A [T]transgenic plant according to claim [21] 20, [characterized in that it is] selected from [the following of] sunflower, tobacco, soya, oilseed rape, "canola", maize, wheat, barley, rice, bean, cassava and peanut.

[23.-] 22. Use of a transgenic plant[s] according to any of [the] claims [21] 20 to [22] 21 for the production of substances resulting from the expression of chimeric genes.

[24.-] 23. Use of a transgenic plant[s] according to claim 22 [23 characterized in that] wherein the substances are proteins, bioactive substances and oils.

[25.-] 24. Substances obtained according to any of claims 23 and 24.

AMENDMENT UNDER ARTICLE 19 PCT

The following pages 28-29 comprise a translation into English of claims 1-25 filed with the International Bureau under Article 19 PCT

CLAIMS

1.- Nucleotide sequence constituted by the *Ha ds10 G1* gene, its promoter, Ha ds10 G1 5'- and 3' flanking sequences and uses thereof in specific gene expression in seeds.

2.- Nucleotide sequence according to claim 1, characterized in that it is comprised of the SEQ ID No. 1.

3.- Nucleotide sequence according to claim 1, characterized in that it is comprised of fragments of SEQ ID Nº1.

4.- Nucleotide sequence, characterized in that it comprises a nucleotide sequence according to any of claims 1 to 3.

5.- Nucleotide sequence homologous to the nucleotide sequence according to any of the claims 1 to 4.

6.- Homologous nucleotide sequence according to claim 5, characterized in that the homology is from 70% to 95%.

7.- Nucleotide sequence characterized in that it contains sequences according to any of the claims 1 to 6 and a chimeric gene.

8.- Nucleotide sequence, according to claim 7, characterized in that it permits the expression of a chimeric gene.

9.- Nucleotide sequence, according to claim 8, characterized in that it permits the expression of the chimeric gene specifically of seeds from early maturation stages.

10.- Nucleotide sequence, according to claim 9, characterized in that it is constituted by the constructions ds10F1, ds10F2, ds102Δ, ds10F3 and ds10EC1 or part of said sequences.

11.- Nucleotide sequence, according to claim 9, characterized in that it contains *Ha ds10 G1* gene coding and 3'-flanking sequences.

12.- Nucleotide sequence, according to claim 11, characterized in that it contains in the constructions ds10F2 and ds10F2Δ.

13.- Nucleotide sequence, according to claim 9, characterized in that it contains *Ha ds10 G1* gene coding and intron sequences.

14.- Nucleotide sequence, according to claim 13, characterized in that it is contained in the constructions ds10F3.

15.- Expression cassette characterized in that it contains a nucleotide sequence according to any of the claims 1 to 14 and a chimeric gene.

16.- Vector characterized in that it contains an expression cassette

according to claim 15.

17.- Host cells characterized in that it contains a nucleotide sequence according to any of the claims 14 to 16.

18.- Use of nucleotide sequences according to any of the claims 1 to 16 in the specific expression of chimeric genes in seeds, seed parts, seed extract, seed embryos and seedling tissues.

19.- Use of the nucleotide sequences according to any of the claims 10 to 12 in order to increase the expression of chimeric genes specifically in transgenic plant seeds.

20.- Use of nucleotide sequences according to any of the claims 12 to 14 in order to increase the expression of chimeric genes in seeds and/or reduce it in other tissues.

21.- Transgenic plants characterized in that they are plants transformed by a nucleotide sequence according to any of claims 1 to 16.

22.- Transgenic plant according to claim 21, characterized in that it is selected from the following of sunflower, tobacco, soya, oilseed rape, "canola", maize, wheat, barley, rice, bean, cassava and peanut.

23.- Use of transgenic plants according to any of the claims 21 to 22 for the production of substances resulting from the expression of chimeric genes.

24.- Use of transgenic plants according to claim 23 characterized in that the substances are proteins, bioactive substances and oils.

25.- Substances obtained according to any of claims 23 and 24.

UNGRIA®

TRANSLATION TO ENGLISH OF INTERNATIONAL APPLICATION AS FILED FOR CHAPTER II

TITLE

PROMOTER AND REGULATOR SEQUENCES *Ha ds10 G1*: A GENE LEA OF SUNFLOWER EXPRESSED EXCLUSIVELY IN SEEDS FROM THE MATURATION PHASE.

TECHNICAL SECTOR

Agriculture. The subject of this invention is related to obtaining of regulatory ("promoter") DNA sequences and the construction of new chimeric genes, using these sequences, capable of being specifically expressed in transgenic plant seeds. *Ha ds10 G1* gene has the peculiarity of only being expressed in sunflower seeds from the maturation until the desiccation phase, without responding to hormones such as abscicic acid (ABA) or water stress in vegetative tissues. Furthermore, gene *Ha ds10 G1* is expressed homogeneously in immature embryos and preferentially in the palisade parenchyma of mature embryo cotyledons. These expression patterns, as well as the high activity levels of the gene, suggest that its regulatory sequences are particularly appropriate for the genetic manipulation of storage substances in seeds.

PRIOR ART

Up to now in order to confer specific expression in transgenic plant seeds, promoters have been isolated, characterised and used, especially belonging to plant genes which code for storage proteins or other products solely expressed in seeds during different phases of development [see the following bibliographical references and patents, as well as other documents cited in them: Thomas TL, in *Plant Cell*, vol 5, pp 1401-1410, 1993; Gatehouse JA and Shirsat AH in *Control of Plant Gene Expression*, pp 357-375, *CRC press*, 1993; and the USA patents numbers: 5530192, 5530194 and 5420034]. For example, this has allowed the obtaining of new transgenic plants with modified fatty acid and storage protein content [see: Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C, Hawkins DJ, Radke SE and Davies HM, in *Science*, vol. 257, pp.72-74, 1992; and Saalbach I, Pickardt T, Machemehl F, Saalbach G, Schieder O, and Muntz K, in *Molecular and General Genetics* 242: 226-236, 1994]. Other promoters with different tissue specificity in seed and varied temporal expression patterns could be useful for the development of the enormous potential of this technique. Recently in our group, and in other laboratories, we described the expression in

seeds of genes that code for low molecular weight heat shock proteins (sHSPs: *small heat-shock proteins*). One of these genes, *Ha hsp17.7 G4*, shows in tobacco transgenic plants, expression patterns appropriate for its possible use in the genetically engineered modification of seeds: this gene is expressed from early seed maturation phases, and is cotyledon tissue specific [Coca MA, Almoguera C, Thomas TL and Jordano J, in: *Plant Molecular Biology* 31: 863-876, 1996]. However, gene *Ha hsp17.7 G4* , like other sHSP plant genes expressed in seeds, is also expressed in response to heat (heat shock) in plant vegetative tissues after seed germination. The latter makes its use in genetic engineering impossible when regulatory DNA sequences that guarantee the absence of expression of chimeric genes outside of the seed are required: for example, when the expression elsewhere of these genes may affect viability, growth or the health of the transgenic plants. To solve these problems we modified the *Ha hsp17.7 G4* gene regulatory sequences such that the chimeric genes that contain these sequences maintain their expression in seeds and lose their heat induction; a procedure which can be used for the modification and similar use of regulatory sequences of other sHSP genes expressed in seed [Almoguera, Prieto-Dapena and Jordano, patent request #9602746 (Spanish Patent Office)]. Alternatively, we have also proposed a similar use for the promoter and regulatory sequences of the sunflower gene *Ha hsp17.6 G1*, that is only expressed in seeds. This gene does not respond to heat or other types of stress (cold, dehydration, ABA hormone treatment) in vegetative tissues [Carranco, Almoguera and Jordano, patent request #9701215 (Spanish Patent Office).

In this application we propose alternative analogous uses for promoter and regulatory sequences of sunflower LEA *Ha ds10 G1* gene. Gene *Ha ds10 G1* has been found in a genomic clone corresponding to a previously described cDNA (*Ha ds10*, access number X506999) whose expression patterns were not totally known [Almoguera and Jordano, *Plant Mol. Biol.* 19:781-792, 1992]. The promoter and regulatory sequences of this gene (*Ha ds10 G1*) have been cloned and are described, characterised and used for the first time in the examples in this application. The *Ha ds10 G1* gene belong to the Class I LEA (Late Embryogenesis Abundant) gene family (D-19 or LEA-I type) These genes code for highly conserved proteins in various plant species, and their expression is usually restricted to seeds and early germination phases [see for example the

following reviews: Dure III, L., *Structural motifs in Lea proteins*, in *Plant Responses to Plant Dehydration During Environmental Stress.*, Close TJ and Bray EA Eds., *Current Topics in Plant Physiology 10*: 91-103, 1993; and Delseny M, Gaubier P, Hull G, Saez-Vasquez J, Gallois P, Raynal M, Cooke R, Grellet F., *Nuclear Genes expressed during seed desiccation: relationship with responses to stress* , in Stress-induced Gene Expression in Plants (Basra, A. S., ed.), pp. 25-59, Harwood Academic Publishers, Reading, 1994]. LEA gene promoters have not been considered as good candidates for their use in seed storage substance modification projects, as in general their activity is expressed in later seed maturation phases, such as embryo desiccation [see the considerations of Kridls JC, Knauf VC, Thompson Ga in *Control of Plant Gene Expression*. pp. 481-498, CRC press, 1993]. However, LEA genes that are activated in maturation phases prior to desiccation are known, such as the cotton genes denominated LEA-A [Hughes DW and Galau GA, *The Plant Cell* 3:605-618, 1991]. Examples of activation prior to desiccation are also known with the class I LEA genes, such as in the case of *At Em1, emb564* and *emb1* genes [in arabidopsis, maize and carrot, respectively: Gaubier P, Raynal M, Hull G, Huestis GM, Grellet F, Arenas C, Pages M, and Delseny M, *Mol. Gen. Genet.*, 238: 409-418, 1993; Williams B, and Tsang A, *Plant Mol. Biol.*, 16: 919-923, 1991; Wurtele ES, Wang H, Durgerian S, Nikolau BJ, and Ulrich TH. *Plant Physiol.* 102:303-312, 1993]. These examples seem to indicate the possible use of regulatory sequences from genes in this family for the modification of seeds. However, its specific use would be limited both by the expression levels obtained in each case and in each development phase; as well as the different tissue specificities. Thus, even though in *Arabidopsis* the *At Em1* gene is activated early, its expression is basically restricted to cotyledon provascular tissue and cortical tissue external to the embryonic axis [Gaubier, P., Raynal, M., Hull, G., Huestis, GM., Grellet, F., Arenas, C., Pages, M., and Delseny, M., *Mol. Gen. Genet.*, 238: 409-418, 1993]. In the case of the carrot gene, *emb1*, its mRNA are preferentially localised in the embryonic meristems, especially in the *procambium* [Wurtele ES, Wang H, Durgerian S, Nikolau BJ, and Ulrich TH. *Plant Physiol.* 102:303-312, 1993]. No gene sequence of the *emb564* gene has been published and the exact localisation of its mRNA is unknown [Williams B and Tsang A, *Plant Mol. Biol.*, 16: 919-923, 1991].

The expression of sunflower gene *Ha ds10 G1*, as well as its promoter and

regulatory sequences present unique characteristics among the other members of the LEA-I family, as described below, which means that these sequences may be potentially used for the modification of seeds by genetic engineering.

DESCRIPTION OF THE INVENTION

In this invention we isolate and characterise in transgenic tobacco plants, the promoter and regulatory sequences of a sunflower LEA-I gene, *Ha ds10 G1*. These sequences (Example 1) present highly appropriate characteristics for their use in the modification of seeds (e.g. storage substances). The advantages of their possible use in transgenic plants are demonstrated through other examples: A.- Studies of *HA ds10* mRNA accumulation and localisation in the homologous system (Example 2). These studies demonstrate both the high expression levels reached during embryogenesis from early maturation phases, as well as the absolute seed specific localization, accompanied of a homogenous distribution in embryos which terminates essentially restricted to the cotyledon palisade parenchyma, a tissue specialised in the accumulation of sunflower storage substances. B.- In example 3, we also illustrate the possible use of such sequences via the construction and analysis of various chimeric genes in transgenic plants, using the promoter and combinations of various *Ha ds10 G1* regulatory sequences (5'-flanking, coding, intron and 3'-flanking), with the reporter gene of bacterial β-glucuronidase (GUS). These examples demonstrate in a heterologous model (tobacco) the usefulness of the different chimeric genes tested: high expression level and specificity to seeds from early maturation phases, as well as the functional contribution of the various sequences tested. Via the examples attached we demonstrate that the seed specificity is basically conferred by the promoter and the 5'-flanking sequences of *Ha ds10G1* (including untranscribed and transcribed sequences: such as the 5'-UTR and part of the coding sequence). Additionally, the 3'-flanking sequences increase expression levels in seeds and the intron specifically reduces it in non-embryonic tissues. Given the conservation of the regulation of embryonic gene expression in plant seeds, including LEA-I genes [Thomas TL, in I 5:1401-1410, 1993]; these sequences could be used both in the homologous system (sunflower) as in other heterologous systems of great economic importance (for example oilseed rape, soybean, maize, etc).

The practical embodiment of this invention, represented by the attached

examples and figures, uses conventional Molecular Biology, Microbiology, recombinant DNA and transgenic plant production techniques that are common practice in laboratories specialised in these fields. These techniques have been explained in sufficient detail in the scientific literature [for example see: Sambrok J, Fritsch EF, and Maniatis T, *Molecular Cloning: A Laboratory Manual, Cold Spring Harbor laboratory Press*, 2nd Edition, 1989; Glover DM, *DNA Cloning, IRL Press* , 1985; Lindsey K., *Plant Tissue Culture Manual, Kluwer Academic Publishers*, 1993; and Gelvin SB, Schilperoort RA, Verma DPS, *Plant Molecular Biology Manual, Kluwer Academic Publishers* , 1992]. For more specific details, the pertinent bibliographical references are cited in the corresponding section in this application.

EXAMPLE 1: Cloning, determination of restriction map, nucleotide sequence and analysis of the *Ha ds10 G1* promoter.

To obtain the *Ha ds10 G1* clone the sunflower genomic DNA gene library described by Coca *et al.* [*Plant Mol. Biol.* 31: 863-876, 1996] was screened, with the probe corresponding to total *Ha ds10* cDNA [Almoguera and Jordano, Plant Mol. Biol. 19: 781-792, 1992]; using standard hybridisation conditions and molecular cloning procedures described in detail in the first of the references (Coca *et al.*, 1996). We thus isolated a phage (IGEM11) with a sunflower genomic DNA insert of approximately 16.5 Kb whose partial map is shown in Figure 1. We determined, using restriction analysis, that the two fragments adjacent to the Sac I site (4.2 and 9.3 Kb) contain the sequences that hybridise with the cDNA. A detailed restriction map of the first of these fragments was determined and part (4 Kb) of the second (Figure 1). Different genomic DNA subfragments, corresponding to the mapped region, were cloned in pBluescript SK+ vector, resulting in plasmids whose names and inserts are listed in Figure 1. The 3617 bp nucleotide sequence between the Sac I and Sma I sites (Figure 1, lower section) was determined from these plasmids on both DNA strands using the Sanger (dideoxy) method. These data are presented in SEQ No. 1. We confirmed by comparing the sequences, that part of the genomic sequence determined corresponds to Ha ds10 cDNA [Almoguera and Jordano, *Plant Mol. Biol.* 19: 781-792, 1992, GenBank access number X59699]. The amino acid sequence of the protein coded by the *Ha ds10 G1* gene is indicated below the corresponding nucleotide sequences. In the genomic DNA, the coding region is interrupted by an abnormally long intron (1024 bp), even though it is situated in a

conserved position in other class I LEA genes [see data reviewed by Simpson GC, Leader DJ, Brown JWS and Franklin T, in *Characteristics of Plant pre-mRNA Introns and Transposable Elements, Plant Mol. Biol. LabFax*, pp 183-252; Croy RRD Ed., Bios Scientific Publishers Ltd. 1993]. The only difference between the gene sequences coding for mRNA and those of cDNA, was a two nucleotide inversion (GC instead of CG) within the second exon (in positions +1176 and +1177 from the initiation codon) which induces an amino acid change (S instead of T) in the protein sequence. The difference is due to an error (due to a compression) in the initial reading of the cDNA sequence reactions. The *Ha ds10 G1* sequences we have determined also include 1576 bp of the gene promoter and 5'-flanking region, and 553 bp of 3'-flanking genomic regions not present in the original cDNA.

Three possible transcription initiation sites were determined in the *Ha ds10 G1* promoter by the primer extension technique. Two of these sites have been confirmed with other techniques (sites 1 and 2, indicated by arrows in SEQ No. 1). For this the procedure described by Domon *et al.* was used [Domon C, Evrard JL, Pillay DTN and Steinmetz A. *Mol. Gen. Genet.* 229:238-244, 1991], total sunflower embryo RNA was hybridised with the synthetic primer: 5'-CTCCTGTTCCGGAATTTTGCGTGT-3', whose sequence corresponds to that of the non coding strand of *Ha ds10 G1*, between positions +25 and +48, from the initiation codon. The hybridisations with the primer were carried out at 62°C. The hybrids were extended with AMV reverse transcriptase, for 90 min at 42°C. The extension products were analysed on 6% PAGE sequencing gels, along with sequence reactions produced using the same primer. Initiation sites 1 and 2 (at positions -33 and -25, see SEQ No. 1) are functional, and are detected independently using the ribonuclease A protection technique (RNAse A, see Figure 3A). A third initiation site (site 3, in position -119 in SEQ No. 1) could not be clearly confirmed with this technique. These initiation sites functionally define the 3' end of the *Ha ds10 G1* gene promoter.

The analysis of the proximal sequences of the *Ha ds10 G1* gene promoter demonstrated that the two initiation sites detected (sites 1 and 2) are found at an appropriate distance from a possible TATA sequence (at position -86). The possible more distal site (site 3, -119) does not have clear TATA sequences in its proximity. Apart from these promoter elements, two possible RY "boxes" (RY1 and RY2 at positions -129 and -65 of SEQ No. 1) were observed, analogous to

those that participate in the regulation of the expression of numerous plant genes in seeds [Dickinson DC, Evan RP, and Nielsen RC, in *Nucleic Acids Research* 16: 371, 1988].

We modified the RY1 box sited at position -129, verifying by transient expression experiments in sunflower embryos, its functional requirement for the trans-activation of the *Ha ds10 G1* promoter by ABI3 type transcription factors [Giraudat J., Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM in *The Plant Cell*: 1251-1261, 1992]. In order to do this we prepared modifications of the ds10::GUS fusions constructed for transgenic plant studies (see Example 6.3 and Figure 5). The chimeric genes contained in these two fusions (ds10F1 and ds10F2) are purified as DNA fragments which were subcloned by ligation into pBluescript SK+ (Promega) vector, thus changing the binary vector sequences for smaller ones, more useful for transient expression experiments. We thus obtained the plasmid pSKds10F1 using the Sal I - Eco RI fragment (with the chimeric gene obtained from ds10F1). In the case of ds10F2, the Sph I - Eco RI fragment (from position -125 in *Ha ds10 G1*, to the 3' end of *nos*) was ligated to the complementary fragment (which contains the promoter and 5'-flanking sequences of *Ha ds10 G1*), purified after digestion of pSKds10F1 with Sph I and Eco RI, resulting in the pSKds10F2 plasmid. Finally, from the pSKds10F1 and pSKds10F2 plasmids (maps not shown) mutagenised versions were obtained by digestion of their DNA with Sph I, blunting the resulting ends by treatment with T4 DNA polymerase, followed by re-ligation of the DNA. We thus obtained plasmids pSKds10F1ΔRY and pSKds10F2ΔRY (maps not shown). These plasmids only differ by a 5 nucleotide deletion between positions -126 and -122 of the *Ha ds10 G1* promoter. These changes destroyed the RY1 box present in the ds10F1 and ds10F2 chimeric genes (see Figures 1, 2 and 5), this was verified by the Sanger (dideoxy) method sequencing reactions, using the primer 5'CTCCTGTTCCGGAATTTTGCGTGT3' (non coding strand of *Ha ds10G1* between positions +25 and +48).

The trans-activation experiments in transient expression were carried out by bombarding sunflower embryos with projectiles coated with DNA mixtures from different plasmids. These mixtures contain a reference plasmid, pDO432 [Ow DW, Wood KV, deLuca M, de Wet JR, Helinski D and Howell SH. Science 234: 856--859, 1996], with the firefly (*Photinus pyralis*) luciferase (LUC) gene regulated by the CaMV 35S promoter, the fusion of ds10::GUS tested in each

case (with intact or modified RY1 sequences), and an effector plasmid, pABI3, which expresses the ABI3 factor under control of the CaMV 35S promoter. pABI3 was obtained by substituting the Pv ALF cDNA from the pALF plasmid [Bobb AJ, Eiben HG, an Bustos MM in *The Plant Journal* 8: 331--343, 1995], with ABI3 cDNA. The ABI3 cDNA was cloned as an Xba I fragment (blunted with Klenow *enzyme*) - Eco RI (partial), fragment purified from the pcabi3-4F plasmid [Giraudat J., Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM in *The Plant Cell* 4: 1251-1261, 1992]. pABI3 plasmid is added to, or omitted from, the mixture to test the effect of the ABI3 factor on GUS expression in the fusion tested. The experiments were essentially carried out as described by Bobb *et al.*, [Bobb AJ, Eiben HG, and Bustos MM in *The Plant Journal* 8: 331--343, 1995], with the following modifications. Sunflower embryos (17-20 dpa) were prepared as follows. Sunflower seeds were sterilised by washing in 70% ethanol for 1 min, and in 2% sodium hypochlorite with a drop of Triton X-100 for 40 min, finally rinsed several times with distilled water, and then peeled under sterile conditions. The embryos are cut longitudinally (separating the two cotyledons) and placed, with the cut surface down on MS solid medium plates, containing 2% sucrose and 0.5 M sorbitol. They are then pre-cultured for 2-4 h in the dark at room temperature (25°C). All the plasmids were purified using the *Quantum midiprep kit* (Biorad). Normally for each bombardment were used: 0.2 μg of reference plasmid, 0.1 μg ds10::GUS plasmid and 1 μg of effector plasmid (or the same amount of pJIT82 plasmid in the negative controls). For the preparation of the gold particles, as well as the DNA precipitation onto them, we followed the method described by Chern *et al.* [Chern MS, Bobb AJ and Bustos M. *The Plant Cell* 8: 305-321, 1996]. The particle bombardment was carried out using the *Biolistic PDS-1000 He* system (Biorad). The bombardment conditions were the following: 1550 psi rupture membrane, 1.6 μm diameter gold particles, distance from rupture membrane to macrocarrier 8 mm, distance from macrocarrier to grid 6 mm, and distance to the tissue to be bombarded 6 cm. The bombarded cotyledons were incubated for 24 h at 28 °C in the dark, after which the GUS activity (relative to LUC activity) was tested as described by Bobb *et al.* [Bobb AJ, Eiben HG, and Bustos MM in *The Plant Journal* 8: 331-343, 1995].

The addition of pABI3 effector plasmid had a clear effect on the relative expression of GUS/LUC in bombarding with the pSKds10F2 fusion (average increase in relative activity ≈ 46.2X). On the other hand, if the trans-activation

was carried out with the same plasmid with a mutation in the RY box (pSKds10F2ΔRY1), a significant reduction in the average increase in relative activity due to the ABI3 effect (≈ 26.3X) was observed. This result, shown in figure 2, confirms the functional requirement of the RY1 sequence (position -129 in SEQ No. 1). Therefore, this RY box participates in the transcriptional activation in seeds of the *Ha ds10 G1* promoter for ABI3 type factors [Giraudat J., Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM in *The Plant Cell* 4: 1251-1261, 1992]. Other promoter sequences (e.g. RY2 in -65) could also contribute to the transactivation effect observed, as the mutation tested does not completely destroy the activator effect of ABI3.

EXAMPLE 2: Accumulation and specific localisation of *Ha ds10* mRNA in sunflower embryos:

The messenger RNA accumulation patterns of the *Ha ds10G1* gene were determined by the Ribonuclease A (RNAse) protection technique , described in detail by Almoguera *et al.* [Almoguera C, Coca MA, Jordano J. *Plant Physiol.* 107: 765-773, 1995]. To do this, total RNA samples prepared from seed embryos at different stages of development under normal growth conditions were used [Almoguera and Jordano, *Plant Mol. Biol.* 19: 781-792, 1992; Coca *et al.*, *Plant Mol. Biol.* 25: 479--492, 1994]; of seedlings 3-day after imbibition (dpi); and of different adult plant organs before flowering. The seedling and plant RNA were prepared from plant material obtained both under controlled growth conditions [Almoguera and Jordano, *Plant Mol. Biol.* 19: 781-792, 1992; Coca MA, Almoguera C, and Jordano J. *Plant Mol. Biol.* 25: 479-492, 1994; Coca MA, Almoguera C, Thomas TL, and Jordano J. *Plant Mol. Biol.* 31: 863--876, 1996], and after stress treatments: water deficit [Almoguera C, Coca MA, and Jordano J. *Plant J.* 4: 947-958, 1993; Coca MA, Almoguera C, Thomas TL, and Jordano J. *Plant Mol. Biol.* 31:863-876, 1996]; or after addition of hormones such as absicic acid [Almoguera C and Jordano J. *Plant Mol. Biol.* 19: 781-792, 1992; Coca MA, Almoguera C, Thomas TL, and Jordano J. *Plant Mol. Biol.* 31: 863-876, 1996]. The conditions used in each treatment are described in detail in the references cited for each case. The riboprobe used to detect the *Ha ds10 G1* mRNA is 396 nucleotide long, of which 63 are sequences of the pBluescript SK+ vector and the rest the sequence of the non coding strand of *Ha ds10 G1* between positions +212 and -121 (Sph I). This hybrid probe with the 5' end of *Ha ds10 G1*

messenger RNAs, exceeding the more distal transcription initiation site (site 3, SEQ No. 1), allows the detection of messenger RNA (mRNA) produced from the three initiation sites and the experimental verification of the initiation positions. This riboprobe was prepared by *in vitro* transcription, using RNA polymerase T3 and as a template ds10G1S3Δ4.4 plasmid DNA (Figure 1) which contains the *Ha ds10G1* sequences between -1576 (Sal I) and +212 cloned in the pBluescript SK+ vector.

The results in Figure 3 show that the *Ha ds10 G1* messenger RNAs are only detected in seeds. Higher accumulation levels are observed around 18-20 dpa, gene expression is detected from 10 dpa and it disappears after germination (Figure 3). Treatments with ABA, or water deficit did not induce the accumulation of *Ha ds10 G1* messenger RNAs (data shown for ABA in seedlings, Figure 3). As a positive control in the RNA samples tested for the different treatments, we carried out hybridisations (data not shown) with another previously described 651 nucleotide riboprobe of *Ha hsp17.7 G4* gene [Coca *et al.*, *Plant Mol. Biol.* 31: 863--876, 1996]; as this gene is expressed in response to the different treatments tested. These analysis showed that the *Ha ds10 G1* mRNAs were only accumulated in seeds, under normal growth conditions and from early stages of maturation, confirming the initiation from at least sites 1 and 2 (indicated in SEQ No. 1). The band marked by the number 3 (Figure 3) does not coincide well with the expected size for initiation site 3 (SEQ No. 1). This band could be due to the protection of messenger RNA sequences of a highly homologous gene, or even *Ha ds10 G1* itself, containing intron sequences (unprocessed mRNA).

The distribution of *Ha ds10 G1* mRNAs in sunflower embryos was investigated by *in situ* hybridisation localisation experiments. In order to do this, embryos were embedded in paraffin, fixed, sectioned and hybridised with specific probes; essentially as described by Molinier [in the thesis: *Diplome d'Etudes Approfondies de Biologie Cellulaire et Moléculaire, Université Louis Pasteur, Strasbourg*, 1995]. The fixing time was increased from 16 h at 4°C to 5 days, the increase depending on the age of the embryos. The dehydration of the fixed embryos was carried out by successive incubations (2 times each for 30-90 min.) in 10%, 20%, 30%, 40%, 60%, 70%, 95% and 100% ethanol; followed by immersion in 100% toluene (1-3h, 2 times). The fixed embryos were first embedded in toluene:paraffin (1:1), at 65°C for 6-15 h, followed by 5 consecutive inclusions in paraffin, at 60°C for 5-15 h. The pre-hybridisations and

hybridisations with the probes were carried out at 45°C. The specific *Ha ds10 G1* riboprobe, corresponding to the mRNA 3' end was prepared as follows: The ds10G1S1 plasmid (Figure 1) was used as a template to prepare two *in vitro* transcription probes [Almoguera C, Coca MA and Jordano J. *Plant Physiol.* 107: 765-773, 1995] marked with DIG-UTP. The ds10-3'(-) is obtained by digesting plasmid DNA with Pvu II and carrying out the transcription with RNA polymerase T3. This probe corresponds to the non-coding strand of *Ha ds10 G1* between positions +1202 (Pvu II in the second exon) and +1592 (3' end). The second probe [ds10-3' (+), used as a control], was prepared digesting *Ha ds10 G1S1* DNA with Bam HI (in the *polylinker*); and carrying out the transcription with RNA polymerase T7. Probe ds10-3'(+) contains the coding chain of *Ha ds10 G1*, between position +870 and +1592. The specificity of the hybridisation was determined by Southern blot experiments similar to those described by Almoguera and Jordano [*Plant Mol. Biol.* 19: 781-792, 1992]. While the hybridisation with a total cDNA probe detects bands corresponding to some 4-5 different genes in the sunflower genome [Almoguera C, and Jordano J. *Plant Mol. Biol.* 19: 781--792, 1992]; using probe ds10-3'(-) we can detect a single gene (with a slight cross hybridisation with another one, data not shown).

The results obtained in the RNA localisation experiments are shown in Figure 4. Probe ds10-3'(-) is complementary and has opposite polarity to *Ha ds10 G1* mRNA, which allows its detection. The results obtained agree with the protection data shown in Figure 3, and demonstrate its accumulation in embryos from 12-15 dpa (Figure 4A) to 21-28 dpa (Figures 4C, F and H). This accumulation takes place to high levels, which can be deduced from the short time required for its histochemical detection (2-4 hours). In immature embryos (Figure 4A) the distribution of Ha ds10 G1 mRNA is homogeneous and comparable (Figure 4B) to that of 18S rRNA , which is detected using another riboprobe corresponding to fragment G (Eco RI) of the radish 18S gene [described by Delcasso-Tremousaygue D, Grellet F, Panabieres F, Ananiev E D, and Delseny, M. in *Eur. J. Biochem.* 172: 767-776, 1988]. In more mature embryos (21 dpa, Figure 4C) the *Ha ds10 G1* mRNA are also localised fairly homogeneously, with a more intense accumulation detected in the vascular bundles (*procambium*), something which is not observed with the 18S rRNA probe nor in this or other development stages (Figures 4D, B and G). Finally, at 28 dpa the *Ha ds10 G1* mRNA are preferentially localised in the palisade

parenchyma, a tissue specialised in the accumulation of storage substances, located in the internal face of cotyledons (Figures 4F and H). The localisations with probe ds10-3' (+), with the same polarity as the *Ha ds10 G1* mRNA, did not give any hybridisation signal, which was a control for the previously described experiments (compare Figures 4C and E). These experiments demonstrated that the *Ha ds10 G1* mRNA expression patterns in sunflower are very special. The expression observed in seeds, with high levels of accumulation from early embryonic maturation stages (10-12dpa), are combined with spatial distributions which change from homogeneity to a greater abundance in storage substance deposit tissues (palisade parenchyma). The distribution and accumulation pattern of *Ha ds10 G1* mRNA is different from that presented by other plant genes belonging to the same family [Wurtele ES, Wang HQ, Durgerian S, Nikolau BJ and Ulrich TH. *Plant Physiol.* 102: 303-312, 1993; Gaubier, P., Raynal, M., Hull, G., Huestis, GM., Grellet, F., Arenas, C., Pages, M., and Delseny, M., *Mol. Gen. Genet.*, 238: 409-418, 1993]. These results indicate the potential usefulness of chimeric genes that incorporate *Ha ds10 G1* regulatory sequences for the modification of seeds by genetic engineering.

EXAMPLE 3: Construction of ds10G1::GUS chimeric genes and their analysis in tobacco transgenic plants:.

As an example of the possible uses of the promoter and the regulatory sequences of *Ha ds10 G1* gene in the construction of chimeric genes with specific expression in transgenic plant seeds, we describe below the construction and analysis of 4 ds10G1::GUS translational fusions in tobacco transgenic plants (Figure 5). These fusions contain the promoter and different combinations of flanking and intragenic sequences of *Ha ds10 G1* gene for its functional analysis. These 4 fusions provide high levels of expression of the reporter gene (GUS) in seeds from early maturation stages (Figure 6), confirming our observations in the homologous system (Example 2, Figures 1-4).

The first of these constructions, ds10F1 (Figure 5) was obtained from the ds10G1S3 plasmid (Figure 1), which contains the genomic sequences of *Ha ds10 G1* between Sal I (-1576) and Eco RI (+1086), subcloned into the corresponding restriction sites of the pBluescript SK+ vector (Promega). The *Ha ds10 G1* sequences between Eco RI (+1086) and position +98 (in the first exon) were deleted by treating with Exonuclease III the ds10G1S3 DNA (previously digested

with Hind III and Pst I), resulting in ds10G1S3Δ10.5 plasmid (Figure 1). This plasmid was digested with Bam HI (*polylinker* restriction target situated immediately adjacent to position +98 of *Ha ds10 G1*), then filling in the digested DNA ends using the Klenow fragment of DNA polymerase I. The DNA was then digested with Sal I, and the 1679 bp fragment containing the *Ha ds10 G1* sequences between Sal I (-1576) and the filled end of Bam HI was purified. This fragment was cloned between the Sal I and Sma I sites of the pBI 101.2 binary vector, resulting in ds10F1, a translational fusion which contains 1576 nucleotides of 5'-flanking *Ha ds10 G1* sequences (from ATG) and the first 98 nucleotides of the coding gene, in phase with the GUS gene (Figure 5). The ds10F2 fusion was derived from ds10F1 by the insertion of a genomic DNA fragment of Ha ds10G1 comprised between positions (Figure 1) +1205 (Pvu II) and Eco RI (≈+4670) . This fragment contains part of the second exon and ≈3370 nucleotides of 3'-flanking sequences (from the termination codon in position +1301); and replaces the nos-3' sequences in the ds10F1 fusion. The Pvu II- Eco RI insert was purified from ds10G1S2 plasmid DNA. For the insertion of this fragment, the ds10F1 DNA was digested with Sac I and the DNA ends were blunted by treating with T4 DNA polymerase I. Then, the DNA thus treated was digested with Eco RI, and the fragment including the *Ha ds10G1* sequences was purified. This fragment was ligated to the previously described Pvu II- Eco RI insert (with the *Ha ds10 G1* 3'-flanking sequences), resulting in the ds10F2 fusion (Figure 4). The ds10F2Δ fusion (Figure 4) was obtained from ds10F2, by the deletion of the *Ha ds10G1* 3'-flanking sequences between Xba I (≈+2830) and Eco RI (≈+4670). To do this, ds10F2 DNA was digested with both enzymes, religating after blunting the resulting DNA ends with the Klenow fragment of DNA polymerase I. Finally, the fourth fusion (ds10F3, Figure 5) was obtained from a Ha ds10 G1 genomic DNA fragment between Sal I (-1576) and Pvu II (+1204), purified from ds10G1S6 plasmid (Figure 1) after digestion with both restriction enzymes. This fragment was ligated with vector pBI101.3 vector, previously digested with Sal I and Sma I. The ds10F3 fusion thus contains the promoter and the same 5'-flanking sequences of *Ha ds10 G1* present in ds10F1 fusion, as well as the first exon (From +1 to +145), the total intron (from +146 to +1169) and part of the second exon of *Ha ds10 G1* (from +1170 to +1204), fused in phase with the pBI 101.3 GUS gene. In all cases the nucleotide sequence corresponding to the fusion zone, between the GUS and the *Ha ds10 G1* sequences, was tested by

sequencing reactions with the Sanger (dideoxy) method, using GUS sequences as the primer: 5'-ACGCGCTTTCCCACCAACGCTG-3'.

The T-DNA in ds10F1, ds10F2, ds10F2Δ and ds10F3 fusions (Figure 5) was mobilised from *A. tumefaciens* (LBA 4404), obtaining different tobacco transgenic plants with independent integrations of each chimeric gene. These plants were obtained and characterised by standard techniques as described in detail by Coca MA, Almoguera C, Thomas TL and Jordano J, [in *Plant Molecular Biology* , 31: 863-876, 1996]. The expression of *GUS* gene was analysed both in developing seeds and under normal growth conditions (without exogenous stress); as in seedling tissues, in the latter case the expression changes induced by ABA and dehydration treatments were studied. The seed analysis were carried out with the original transgenic plants (T0), while those of the seedlings used descendants of these plants (T1), segregating for the chimeric genes. Quantitative studies by fluorometric analysis of GUS expression levels and their temporal patterns, as well as qualitative studies which analysed histochemically the spatial patterns of expression (tissue specificity) were carried out. These studies were carried out as described in detail by Coca MA, Almoguera C, Thomas TL and Jordano J, [in *Plant Molecular Biology*, 31: 863-876, 1996]. In total, the following number (in parenthesis) of tobacco transgenic plants, T0 "functional", containing the chimeric genes ds10F1 (14), ds10F2 (7), ds10F2Δ (8) and F3 (23) were obtained and analysed. These plants showed high levels of GUS gene expression in seeds (as a result of the activity of the *Ha ds10 G1* gene promoter and regulatory sequences), as illustrated in Figure 6 (panels A-C). The integration of the different chimeric genes in the transgenic plants' DNA was characterised by *Southern* analysis using probes for the coding GUS gene region; PCR amplifications of the sequences close to the ds10::GUS splice, using the 5'-ACGCGCTTTCCCACCAACGCTG-3' (GUS) and 5'-GAGTGAACAgAATtcCATCACAACAGGG-3' (ds10Eco RI) primers; or by the Kanamycin resistance segregation test (conferred by the *nptII* gene), performed as described in [Jordano J, Almoguera C, and Thomas TL, *The Plant Cell* 1: 855-866, 1989]. These analysis determined that the T0 plant selected for the seed expression studies contained 1 to 5 integrations independent of the corresponding chimeric gene. Figure 6 (joined to this application) illustrates the more relevant results obtained in the study of the expression of the chimeric genes analysed in transgenic plants. These results are described in detail below.

GUS expression during seed maturation under controlled growth conditions (without exogenous stress), was analysed by fluorometric (Figure 6A) and histochemical (summary in Figures 6B-E) assays. The fluorimetric assays were carried out in seeds at defined maturation stages, 12, 16, 20, 24 and 28 days post-anthesis (dpa). For each T0 plant and maturation stage, two different floral capsule extracts were prepared, and the GUS activity was assayed with Methylumbelliferylglucuronide (MUG) in duplicate (in total four activity determinations per development stage and per individual transgenic plant). The statistical significance of the differences observed with the different GUS fusions was determined, after log normalisation of the data obtained, by variance analysis [ANOVA, see: Nap JP, Keizer P, and Jansen R, in *Plant Molecular Biology Reporter* 11: 156-164, 1993]. The histochemical assays were carried out with material dissected from seeds, at defined development stages, from the following number of transgenic plants: d10F1, 5, ds10F2, 6, ds10F2Δ, 6 and dsF3, 19. The endosperm and the embryos dissected from individual seeds were stained with X-gluc, for 150 min at 25°C, approximately 150 seeds from each transgenic plant were analysed in this manner.

All chimeric genes produced high levels of GUS expression in seeds, reaching average maximum values of 1.65×10^6 pmol MU/ mg x min (Figure 6A: at 24 dpa). The histochemical assays confirmed these high activity values, since both the embryos (Figures 6B and C) and the endosperm (Figure 6C) were strongly stained from 12 dpa (Figure 6B) and with only 150 min of reaction. In both cases fairly homogeneous spatial distributions of the GUS activity were observed (Figure 6B-C). Furthermore, these expression patterns do not differ qualitatively between the different chimeric gene transgenic plants (data not shown).

The fluorimetric assays revealed interesting quantitative differences between the different ds10::GUS fusions. These differences depend on the *Ha ds10 G1* sequences present in the fusions. In some cases the statistical significance of these differences could be demonstrated (with a confidence level of 95%), which experimentally demonstrates the contribution of the different sequences tested (promoter and 5'-flanking sequences, coding sequences, 3'-flanking and intron) to the embryonic expression patterns observed. The presence of *Ha ds10 G1* 3'-flanking sequences in the fusions increases the GUS expression levels in seeds between 20 and 28 dpa (compare fusions ds10F2 and

ds10F2Δ, with ds10F1 in Figures 5 and 6A). This difference is statistically significant (for example at 28 dpa: F = 5.397, P: 0.0213), and is caused by the *Ha ds10 G1* sequences present in the ds10F2Δ fusion (see Figure 5); since no significant differences were found between the GUS activity of ds10F2 and ds10F2Δ (for example, also at 28 dpa, F=0.274, P=0.6015; see Figure 6A). In the case of ds10F2Δ, the stimulating effect of the 3'-flanking sequences also occurs and is highly significant, in earlier embryonic maturation development stages (Figure 6A, 16 dpa; F=16.607, P=0.001). On the other hand, in these stages (between 12 and 16 dpa) ds10F1 and ds10F2 GUS activities do not differ significantly (e.g. at 16 dpa: F=2.762, P=0.0983; see Figure 6A). Overall these results show that ds10F2Δ is the constructed and tested fusion that works the best in tobacco seeds from 16dpa; and that this is due to the effect of *Ha ds10 G1* 3'-flanking sequences included in it. We do not know if this effect is caused by transcriptional activation or mRNA stabilisation mechanisms, or by a combination of both. In any case the effect is clear and the potential usefulness to design new chimeric genes with more efficient expression in seeds, from relatively early embryonic maturation stages (see also the section "Other Examples").

On the other hand, the comparison of the GUS activities in plants with the ds10F1 and ds10F3 fusions allowed us to investigate the possible effects of the presence of the intron (and/or *Ha ds10 G1* coding sequences in which these fusions differ, Figure 5) on the expression of both fusions. In transgenic tobacco seeds these comparisons demonstrate that the presence of the intron (plus the first total exon and part of the second exon) does not have positive effects on GUS expression, which must be therefore essentially conferred by the *Ha ds10 G1* promoter and the sequences present in ds10F1 (Figure 6A). Thus for example, the activities of ds10F1 and ds10F3 are not statistically different between 12 and 28 dpa, except at 20 dpa (F=4.73, P=0.031) and then the presence of additional sequences in ds10F3 significantly reduced the GUS activity observed. Therefore, even though it is highly probable that the intron is correctly processed in the seeds of heterologous systems such as tobacco (we do not have any formal proof), its possible regulatory role in embryonic development is unclear. However other observations do not exclude that the additional *Ha ds10 G1* sequences in ds10F3 (including the intron) may have regulatory roles in other tissues (see below the effect of these sequences on residual expression of ds10::GUS fusions in pollen and seedlings).

Embryonic specificity (to seeds) of GUS expression conferred by the *Ha ds10 G1* sequences in tobacco transgenic plants was verified through investigations in other tissues; both in the absence of stress as well as after dehydration and ABA treatments. In the case of T0 plants, the only tissue where GUS activity was detected by fluorimetric and histochemical assays, was mature pollen. In other tissues the activities detected barely exceeded background levels (non-transformed tobacco plants). For example, in T0 plant leaves of about two months of age: 0-50 pmol MU/ mg x min. The activities detected in pollen are marginal (almost three orders of magnitude less) when compared with those of seeds from the same transgenic plants. Furthermore, this expression could be an artefact and depend on the use of GUS gene as an indicator in the fusions [according to Uknes S, Dincher S, Friedrich L, Negrotto D, Williams S, Thompson-Taylor H, Potter S, Ward E, and Ryals J, in *the Plant Cell* 5: 159-169, 1993]. However, surprisingly we observed that the activity measured in the pollen of the 9 ds10F3 plants was (136 ±64 pmol MU/ mg x min) significantly less than that of the 5 ds10F1 plants (6427 ±1294 pmol MU/ mg x min; F= 72.573, P= 0.0001). The latter could indicate that, unlike what is observed in seeds during most of their embryonic development (Figure 6A), the presence of the additional *Ha ds10 G1* sequences in ds10F3 (including the intron) may reduce the expression of the chimeric genes containing them in other tissues or stages of development.

The possibility of expression of the ds10::GUS fusions being induced by hormones (ABA) or stress treatments (water deficit) in tobacco transgenic plants (T1) at different times in its vegetative cycle was also checked. In order to do this, we selected descendants of 8 different original plants, after germination in MS medium with 300 µg/ml kanamycin, containing ds10F1, ds10F2Δ and ds10F3; and another 6 with ds10F2. The resistant seedlings were transplanted in MS medium. Various experiments were carried out with seedlings, both at 8 and 15 days after imbibition. For the ABA treatments, the seedlings were transplanted in MS plates supplemented with 100 µM ABA and cultivated in this medium for 4 days at 25°C in light. The control seedlings were also transplanted in MS medium without ABA. Water stress was induced by placing the seedlings for about 5-6 hours in a laminar flow hood between two filter papers. After the different treatments, the seedlings were processed either individually (for the histochemical assays with X-gluc, by 14 h incubations at 25 °C); or jointly (pool

analysis), for the GUS activity fluorimetric assays as described previously. The adult transgenic plant treatments, were carried out using individual plants propagated as vegetative clones obtained from each original plant. To do this, the seedlings selected from each transgenic plant were transplanted to vermiculite imbibed with Hoagland 0.5X medium. From each seedling three complete explants were obtained, which were placed in hydroponic culture, after recovery, in liquid Hoagland medium (0.5X). The experiments were carried out when the plants had completely recovered from the propagation process, and had roots, stem and about 10-12 leaves. Therefore, genetically identical plants from each selected transgenic seedling were used for the different treatments. The ABA treatments were carried out by adding the hormone to the medium (100 μM) and analysing the GUS activity in the plants after 24h. Water stress was induced by removing the root from the container with the medium, also analysing the plants 24h after starting the treatment. The effect of the different treatments was assessed in three independent experiments performed with the following number of T1 plants for each fusion (the number of T0 plants from which they proceed in each case is given in parenthesis): ds10F1, 11 (6); ds10F2, 10 (5); ds10F2Δ, 5 (3); and ds10F3, 10 (5).

The experiments carried out both in seedlings and in adult plants which confirmed the embryonic specificity of the expression conferred by the *Ha ds10 G1* sequences to the different fusions, also providing additional clues to the possible regulatory role of the *Ha ds10 G1* sequences present in ds10F3 (including the intron) previously mentioned. Thus, both in control adult plants as in treated plants minimum GUS activities (from 3 to 300 pmol MU/ mg x min) were detected in all the tissues analysed (roots, stem, leaves and apical meristem). These activity levels are only slightly above the background levels and can only be detected fluorimetrically (data not shown).

In 8 dpi seedlings the expression of all the fusions is about two order of magnitudes lower than the maximum levels reached in seeds. This expression rapidly decreases between 8 and 15 dpi (e.g. ds10F1 goes from 2864 ±182 to 813±104 pmol MU/ mg x min); and is exclusively restricted to embryonic tissue (cotyledons), without it being detected in other vegetative tissues (radicle, hypocotyl, leaves) differentiated after germination (Figures 6D and E, and data not shown for the other fusions). These results confirm in transgenic tobacco plants the embryonic specificity of the regulation by *Ha ds10 G1* sequences.

Apart from the general reduction in GUS activity values mentioned previously, differences between the values of the different fusions, some statistically significant, were observed. These differences were qualitatively similar to those observed in seeds (Figure 6A). Among them, and for its possible applied interest, we illustrate the reduction of expression after germination, mediated by the *Ha ds10 G1* sequences present in ds10F3 (including the intron). This effect is observed as a significant reduction of GUS activity when the ds10F1 and ds10F3 plant expression patterns are compared (Figures 6D and E). The statistical analysis of the quantitative ds10F1 and ds10F3 data confirmed the significance of this difference, both at 8 dpi (F= 4.36, P= 0.04) and at 15 dpi (F= 4.39, P= 0.039). Additionally, a moderate induction of GUS by ABA treatment in ds10F1 seedlings was observed, which is statistically significant (from 2864 ±182 to 5790 ±733 pmol MU/ mg x min; F= 5.413, P= 0.023). In the case of ds10F3 there was no significant induction by the same treatment (from 1502 ±195 to 2338 ±211 pmol MU/ mg x min; F= 2.58, P= 0.11). The different treatments did not substantially affect the tissue specificity, or the order of magnitude of the expression observed for the different ds10::GUS fusions (data not shown).

OTHER EXAMPLES:

Other chimeric genes can be obtained, in an analogous manner to that of the one described in detail in the previous example, which contain 5'-flanking, and(or) 3'-flanking (terminators), and(or) coding sequences from *Ha ds10 G1*, combined with sequences from other genes. These examples do not involve any additional technical complications to those described in more detail in the previous sections, for which reason they can be easily carried out by persons with sufficient knowledge in the sector of the invention technique. Thus for example, in ds10::GUS fusions the *Ha ds10 G1* could have included other longer 5'-flanking (Figure 1) sequences of the same gene to increase its expression level in seeds as we described in [Coca MA, Almoguera C, Thomas TL, and Jordano J, in *Plant Molecular Biology* , 31: 863-876, 1996]. Equally, the GUS sequences could be substituted by others coding for different proteins or peptides (natural or artificial), whose regulated production in plant seeds could be of industrial interest. Examples of these last possibilities, non exclusively, would be the fusion with *Ha ds10 G1* sequences of coding sequences of genes involved in fatty acid biosynthesis in seeds [Voelker TA, Worrell AC, Anderson L, Bleibaum J, Fan C,

Hawkins DJ, Radke SE and Davies HM, in *Science*, 257:72-74, 1992], of storage proteins with compositions rich in specific amino acids [Saalbach I, Pickardt T, Machemehl F, Saalbach G, Schieder O, and Muntz K, in *Molecular and General Genetics* 242: 226-236, 1994], or peptides with antigenic or pharmacological activities [Vandekerckhove J, Van Damme J, Van Lijsebettens M, Botterman J, De Block M, Vandewiele M, De Clercq, Leemans J Van Montagu, M and Krebbers E, in *BioTechnology* 7: 929-932, 1989]. These fusions would be carried out and used in an analogous manner to what is described in the publications cited as an example (not exclusive) in each case. To facilitate these possibilities, we have created a plasmid (ds10EC1) that contains an expression cassette including the promoter and the 5'- and 3'-flanking sequences of *Ha ds10 G1* present in ds10F2Δ (see Figure 5). Between both sequences and by directed mutagenesis [Chen E and Przybila AE, in *BioTechniques* 17: 657-659, 1994] we have added an Eco RI restriction site, which allows the insertion of gene, or corresponding peptide sequences, as mentioned previously (available in other laboratories, or that could be designed or synthesised). The ds10EC1 plasmid was constructed from ds10G1S3Δ10.5 (Figure 1). From this plasmid, we amplified the *Ha ds10 G1* sequences between positions -1574 (Sal I) and +98 by PCR; using DNA polymerase Pfu and the primers 5'-ATTAACCCTCACTAAAG-3' (T3) and 5'-GAGTGAACAgAATtcCATCACAACAGGG-3' (ds10Eco RI). In the latter the three sequence changes (indicated in lower case letters) introduce the new Eco RI site in the position of the initiation codon. After PCR a 199 pb (*megaprimer*) DNA fragment is purified, which along with the 5'-AATACGACTCACTATAG-3' (T7) primer is used for a second PCR amplification of ds10G1S3Δ10.5. The amplified DNA (795 pb) was digested with Eco RI and Sph I. The resulting DNA fragment (125 pb), with the *Ha ds10 G1* sequences between Sph I (-126) and the new Eco RI site, was purified and ligated; replacing in ds10G1S3 the *Ha ds10 G1* (Figure 1) sequences between positions -126 (Sph I) and 1086 (Eco RI). After this step, the PCR amplified sequence was verified by sequencing (Sanger's method) using the T3 primer. Finally, an Ha ds10 G1 genomic DNA fragment (Figure 1) was inserted in the plasmid obtained in the previous step, with sequences between +1086 (Eco RI) and ≈+3000 (Xba I), obtaining the ds10EC1 *cassette* (Figure 4), cloned in the pBluescript SK+ plasmid. The 3' end of ds10EC1 DNA differs from that of ds10F2Δ only by 119 additional nucleotides, corresponding to the intron and second exon sequences

of *Ha ds10 G1*. Furthermore, the *Ha ds10 G1* sequences in ds10EC1 differ from the corresponding ones in ds10F2Δ in the absence of nucleotides 1-98 of the first exon (Figure 5).

Given that the presence of additional *Ha ds10 G1* sequences in ds10F3 (including the intron, the first exon and part of the second exon) reduced the expression of this chimeric gene specifically in non embryonic tissues (Example 3, Figures 6D-E), it is conceivable that such sequences may be used to confer seed specificity to other chimeric genes with different promoters. The design of such chimeric genes does not involve additional technical difficulties other than those described in the previous sections: see for example the detailed procedures on the use of plant introns to prevent the expression of chimeric genes in *Agrobacterium* [Mankin SL, Allen GC and Thompson WF. *Plant Molecular Biology Reporter* 15: 186-196, 1997]

The chimeric genes containing the *Ha ds10G1* regulatory sequences could be transformed to other plants different from tobacco (the model system used in example 3). Among these there are plants with major economical interest such as: sunflower, soybean, oilseed rape, "canola", maize, wheat, barley, rice, cassava, bean, peanuts, etc. whose genetic transformation is possible and has been sufficiently documented in the scientific literature: see for example Lindsey K, Ed. (1993). [*Plant Tissue Culture Manual*. Kluwer Academic Publishers]; and the review by Christou [*Trends in Plant Science*. 1: 423- 431, 1996]. The results shown in example 3 demonstrate that, in tobacco, the genes constructed with the *Ha ds10 G1* regulatory sequences have a high activity from relatively early embryonic maturation stages, and also maintain the seed specificity characteristic of *Ha ds10 G1* in sunflower. These results could also be obtained with other plants, such as those mentioned previously.

DESCRIPTION OF THE FIGURES:

Figure 1. Upper section: restriction map of the *Ha ds10 G1* genomic sequences flanking its coding region. The continuous lines on the map indicate the different genomic DNA fragments subcloned in pBluescript SK+ vector (the names of the respective fragments are indicated over each fragment). The plasmids prepared by Exo III deletions are indicated over the original plasmid (ds10G1S3ΔSacI), indicating in each case the deletion end. On the lower section of the figure a detailed restriction map of the region whose nucleotide sequence was determined is shown. The extension of the different reactions used to

assemble the different sequences of both DNA strands, are indicated by horizontal arrows (above the map for the coding strand, and underneath the map for the non-coding strand). The transcription initiations sites are indicated with arrows. Scale bars are included for both maps.

Figure 2. Functional implications of the RY1 (-129) sequences in the trans-activation of the *Ha ds10 G1* promoter. Transient expression experiments carried out after bombarding the sunflower embryos with DNA coated micro-projectiles. The results of 5 independent experiments, in which the different plasmid mixtures (described in Example 1) where bombarded five times in each experiment, are presented. The average β-glucuronidase (GUS) activities normalised versus luciferase activity (LUC), as well as the standard error (indicated with bars), are presented. Key: F2, pSKds10F2; F2ΔRY1, pSKds10F2ΔRY1; ABI3; samples with the effector plasmid. A significant decrease in the relative GUS/LUC activity is observed, due to a mutation in the RY1 box. The basal activities for pSKds10F1 (without including the effector plasmid) are of the order of 46±8.

Figure 3. Accumulation patterns of *Ha ds10 G1* gene mRNA in sunflower. The autoradiograph shown corresponds to the RNAse A protection tests, after hybridising a gene riboprobe with different total RNA samples. An accumulation of messenger RNA produced from *Ha ds10 G1* transcription initiation sites (as protected fragments indicated by the numbered arrows) is observed. These fragments are only detected in embryos (Emb) from 10 to 20 dpa and in mature seeds (25 dpa), but not in other samples tested, such as seedlings (Germ) or seedlings treated with ABA (Germ + ABA). The carrier tRNA corresponds to control hybridisations with yeast tRNA. The bands corresponding to the mRNAs produced from the different initiation sites are indicated with numbers and arrows. The initiation site number 3 (indicated in parenthesis) has not been experimentally confirmed by primer extension. On the left margin are included molecular size markers (pBR322/Hpa III).

Figure 4. Localisation of mRNA in sunflower embryos sections at 12 (A and B), 21 (C-E), and 28 dpa (F-H). The following riboprobes were used in each case: ds10 (-), A, C, F, H; ds10 (+),E, and 18S rRNA, B, D, G. Scale bar = 500 µm (Except in F, 125 µm). Palisade parenchyma= pp. The arrows mark the

procambium.

Figure 5. Restriction maps of ds10::GUS fusions and optimised expression ds10 EC1 cassette, constructed in Examples 3 and 4. The *Ha ds10 G1* and other genes contained in each case are shown by shading of varying intensity. The transcription initiation sites from the *Ha ds10 G1* promoter are marked with arrows.

Figure 6. Expression of ds10::GUS fusions in tobacco transgenic plant seeds. Panel A: Summary of quantitative data (fluorimetric determinations). The average of GUS activities in transgenic plant seeds (T0), and its evolution through the different embryonic development stages is shown. The data corresponding to each fusion are indicated by the symbols in the upper left-hand insert. The bars indicate the standard errors. Panels B-E: representative selection with results of the histochemical GUS activity localisation experiments: B.- embryos at 12 dpa (plants ds10F2Δ, T0). C.- embryos and endosperm at 16 dpa (ds10F2Δ plants, T0). D.- seedlings at 15 dpi under control conditions (ds10F1 plants, T1) E.- seedlings at 15 dpi under control conditions (ds10F3 plants, T1) In panels D and E, the arrows indicate the plant tissue without GUS activity (leaves and hypocotyl).

LIST OF SEQUENCES:

SEQ No. 1: *Ha ds10 G1* gene nucleotide sequence. The transcription initiation sites experimentally determined (site 3, which has not been confirmed by primer extension is indicated in parenthesis) are indicated by arrows. The coding zone is shown by its amino acid translation indicated by the letter (LO1 etc.) code underneath the nucleotide sequence. The termination codon is indicated by an asterisk. The sequence is numbered (on the left margin) starting from the initiation codon. The intron sequences are shown in lower case letters. The TATA box (in position -86) and RY box (-129 and -65) mentioned in the text (Example 1) are shown underlined.

LIST OF SEQUENCES

<110> Consejo Superior de Investigaciones Cientificas

<120> Ha ds10G1 promoter and regulatory sequences: a LEA gene expressed only in seeds

<130> Ha ds10G1

<140>
<141>

<160> 1

<170> PatentIn Ver. 2.1

<210> 1
<211> 3617
<212> DNA
<213> Helianthus annuus

<220>
<221> exon
<222> (1577)..(1721)

<220>
<221> exon
<222> (2746)..(2879)

<220>
<221> intron
<222> (1722)..(2745)

<220>
<221> misc_feature
<222> (1448)..(1455)
<223> RY box

<220>
<221> misc_feature
<222> (1512)..(1519)
<223> RY box

<220>
<221> polyA_signal
<222> (3139)..(3144)

<220>
<221> polyA_signal
<222> (3220)..(3225)

<400> 1
gtcgacttct tcatcttcgt ctaagtgttg agtatcgagt acaaattatt catcttcgtt 60
gtcatcgtat atgagaagca tgtgtttaac attatctctt ggatattgag acggtggact 120
ccgataagca aacgggtcaa aagcgttttg ggcttccgta ggatactcgt acacacccgg 180
tgcaaccatc gtgggacttt gtggtaatga aaaaaactgt ggtggatagt tgggttgggg 240
ttgggtgttg taaaaacccg tttgtggaaa aaatgtggtg ggaagttggg ttggtacaca 300
atagatccac ctcgtctact ttgcgagccg cgccccctcg cattcgaacc ggaatccgat 360
acttttttc ttttcatgac ccttgttttg tctatccatg gtatgcatct gaattggtaa 420

attgggtgaa aaatggaatg tttaacatgg taaaatggaa tgttataatt ataaaggtat 480
ttaatgtttt tttttaaacc ataaacggtc atatagccgt ttaaagcaca acggtcaagt 540
ccccaacggt caaatcaagc caaacaatca agtcccgcat gtggcatata tcccctgctt 600
tgaaccaggg ccggcccaag cctaagaatt ttgaggcttg ggctttaggc ctctaaatca 660
atagggactc taacttaaaa aaattatata tgatatttag gttaagttga atttatcttt 720
atatgtacaa aaaaatatat ataattccaa aataaaaaat aaaaaaataa attaacttcg 780
ccaacaataa atcttttgtt atgttttgcc ttttgttttt ttaaaaaaag gctcaaattt 840
ttaatttgct ttaggccacc aaaatggtta aaccgacctt gctttgaacc atacccacac 900
gacaaattag gggatacgga gtgggggtcg gcaaaccaat gccaataaac tttgccgatc 960
aaagtttcac catttcggtg attgtttgcc gatgcggcga ggaggaagaa gagagaggag 1020
ggtagtgtga ggtggggtcc attccaatct caactaatca catttttttc ctttttttta 1080
tttaaaaagt ttaccacttc actaaatgtc taaaccattg ccaacacttt tcaccaaagt 1140
ttaaacattt ttctctgatt gacgtggcac actctcattg gttgattttt tagtttgcca 1200
ctctcaattg tttaaccact ccttacaccc tcttatgtga gcggtggtgt tccccaagcg 1260
acaaaagggc ttaccgcaac cccttaccgc ttccaacctt tacaccctta tcttctgcct 1320
atactgcatg tcactctatt gcagactatc tgagatagct acaacctaac gacttaaggt 1380
gaggacacgt gtatctccaa aacccacttc gtcaccctta ccaccacgtc atcataccac 1440
gtgccaacat gcatgctagt atcttctcta ctatacacat acttatgtac tatatatatt 1500
caccaaaatt acatgcatgt tagcaccctc aattgtacaa cacaacttaa caacttaagt 1560
taaaagagtg aacaaaatgg catcacaaca gggacaacaa acacgcaaaa ttccggaaca 1620
ggagaagaag gatctcgacc aacgagcagc taaaggcgag accgttgttc cgggtggtac 1680
tcgtggcaaa tctcttgagg ctcaagaacg tcttgctgaa ggtatatgct tatttagtta 1740
aatttacata cgtattgggg atggccgttt gatatttagc aaacggcaac ccgaatattt 1800
aggggcccat gatttgacat acatatttga ataagattaa ggggcctaga ggtgtacaaa 1860
tccctttaaa aaagggtccg gttttaaccg gacggaaatg gtttctaaaa tagcggttcg 1920
gtttttaact ggttttacat ggtgggcagt cgaactggta ctagtagatt cggtaagcga 1980
tttgggacaa aaaacgtttt tttaactggt tttgaagatc agtaatacta accggtatgc 2040
aaactgccag ttcgggttga aactggtttc tttttctttt tcaaaaaaga aacggtatgg 2100
gacaaaaaac gttttttttaa actggttttg aagatcaata atactaaccg gtatgcaaac 2160
cgccagttcg gttcgggtta aaactggttt ttttttttt tttaaagaag cggtaaaaaa 2220
accggtttcg gtcataaacc gatttttta acacctacaa gggggccatg atttaaaaaa 2280
aaagcggtaa aaaaaaaccg gtttcggtca taaaccgatt tttttaacac ctacaagggg 2340
gccatgattt aagatgaagt atcggagctc gtgatttaac ataaagtacc tcaaacggcc 2400
gcagtttgat tcaagcgcaa acccgtcgtt ctacgactag tggcgaagct tgagatctcc 2460
gattgggggt cgaaaacgtt tatacccaaa aatttctata gaacgggggt cgaaaacgtg 2520
tatactcaaa aaattctata cgaaagctac atacctgagc gaaaagttcg ggtctggcac 2580
acacccccctc ccccctcttc taacctacgc ccctagctat gacggtgtgg gtagagtcca 2640
ccctttcagt agctttttttg cgaattcaca ttagtttatt ttatagttgt agtgatgcat 2700
aataatatat gcatgtactt aattttgtgt ttggtggtgg tgaagggcgg agcaagggag 2760
gacaaacgag gaaagaccag ctgggaactg aagggtacaa ggagatgggg aaaaagggcg 2820
gtcagaccac cggtgacaag tcggctggtg agcgagagga ggaggaggag gaggactaga 2880
tagtcaatag tggtgtgatg gtgtttgcat gtacgatgat gttaatttcc atgttttata 2940
tatgtgtatg tacctgtagt atggtttagc tcgtgtttca tgttttgttg gtcgttttgg 3000
tatcttcttt agtgcatgta cgactagtag tcctatgatg atgtgatgtg atttgcatat 3060
gttgatataa tggtaagtac aagttttatc tatctatctt ttctttcggc tagagtttac 3120
aaaatgtgta cggatgacaa taataagcaa cttatatttt tccatacaag cttattatgt 3180
gactattgcc caataaggca ataagtaagt gagccaataa ataatagccg gttatgctgt 3240
aacaactatt gtgacaacaa tattctttta ccgaataaat aaaaataagt aagtattagg 3300
agccaataaa taaataaccg gtttaatgag aaaagcacca aaatagcctc tttgatcaat 3360
agaattatga aaatagtctc gactcgacct ttttgaataa cccttcatct cctgtgttgc 3420
gcgatcttcg cccggacgac acgtgtctag tgaacatgcc tattctgatt agtggtgcac 3480
aaatcgggtt ttttttaaa aaacgggttt cgggtaggat cgggtttttt taaaaccggg 3540
tcgggtccgg gtcgggtttt ctaccaaaaa tgtataccct gtataccggg ttcgggtcgg 3600
gttttataaa acccggg 3617

CLAIMS

1.- Nucleotide sequence of the *Ha ds10 G1 sunflower gene,* including its promoter and specific regulatory elements of seeds, described by SEQ Nº 1, and by the restriction maps in Figure 1; and characterized in Examples 1-3.

2.- The sequences, or part of them, identical or homologous to SEQ Nº1 or its complementary sequence (at least by 70%, for example by 80% and particularly less than 95%)

3.- Genes that contain the sequences mentioned in claims 1-2 and that are specifically expressed in seeds, in a homogeneous and abundant manner, from early stages of maturation. These genes may be constructed and used by recombinant DNA techniques, according to the details in the following claims (3-6):

4.- Use in order to confer specific expression in seeds, by means of recombinant DNA techniques, of the *Ha ds10 G* promoter and 5'-flanking and coding sequences *1* (or part of said sequences), contained in the constructions: ds10F1, ds10F2 ds10F2Δ, ds10F3 and ds10EC1 (described in Figure 5).

5.- The use of Ha ds10 G1 coding and 5'-flanking sequences (or part of said sequences), contained in the constructions ds10F2 and ds10FΔ, in order to increase chimeric gene expression specifically in transgenic plant seeds.

6.- The use of coding and intron sequences of *Ha ds10 G1* (or part of said sequences), contained in the construction ds10F3, in order to increase the expression of other chimeric genes in seeds, and/or to reduce it in other tissues, thus increasing the effectiveness and specificity in seeds of these chimeric genes.

7.- Add to the above: seed, part of the seed and seed extract.

8.- Expression cassette that contains a sequence described in claims 1 to 6.

9.- Vector(s) that contains(contain) a sequence described in claims 1 to 7.

10.- Host cells that contain a sequence described in claims 1 to 7.

11.- The process of obtention of transgenic plants characterized in the transformation of a plant (for example, sunflower, soybean, oilseed rape, "canola", maize, wheat, barley, rice, bean, cassava, peanut, tobacco, etc.), with an expression cassette described in claim 8.

12.- Production procedures, for example of coil, proteins or of bioactive substances, by using transgenic plants such as the ones described in claim 11.

13.- Products, for example, oil, proteins or bioactive substances, obtained according to claim 12.